

DIPLOMATURA D'ESTADÍSTICA

Conversió a l'entorn SPAD Windows de dos procediments informàtics aplicats a la inserció de dades d'enquesta

Realitzat per Eduard Solanas Santacana
Dirigit per Roser Rius Carrasco
Juliol de 2003

UNIVERSITAT POLITÈCNICA DE CATALUNYA
Biblioteca



1400503622



Facultat de Matemàtiques
i Estadística

UNIVERSITAT POLITÈCNICA DE CATALUNYA

DADES DEL PROJECTE:

Nom de l'estudiant: EDUARDO SOLANAS SANTACANA

DNI:

Títol del Projecte: CONVERSIÓ A L'ENTORN SPAD WINDOWS DE DOS
PROCEDIMENTS INFORMÀTICS APLICATS A LA INSERCIÓ DE DADOS
D'ENQUESTA

Director del Projecte: ROSER RIUS I CARRASCO

Tutor del Projecte: —

QUALIFICACIÓ

MATRICULA D'HONOR (10,0)

MEMBRES DEL TRIBUNAL (nom i signatura)

President: TOMÁS ALUA I BANET



Vocal: ALBERT AVINYÓ ANDRÉS



Secretari: ROSER RIUS I CARRASCO



Data: 23 DE JULIOL DE 2003

Índex

1.Introducció	1
1.1 Objectius	1
1.2 Les Anàlisis Factorials Descriptives	1
1.3 Procediment per una ACP	2
1.4 Resultats d'una ACP en SPAD	3
1.5 Inserció de dades d'enquesta	4
2.Procediment MIVAR	7
2.1 Plantejament	7
2.2 Disseny del procediment	8
2.2.1 Aportacions al procediment MIVAR	9
2.3 Aplicació procediment MIVAR	9
2.3.1 Anàlisi en components principals de l'enquesta de referència	10
2.3.2 Resultats del procediment MIVAR	11
3. El procediment GREF1	13
3.1 Plantejament	13
3.2 Inserció a partir d'una ACP	14
3.3 Condicions de validesa d'una inserció	15
3.3.1 Coherència global: Test de correlacions	15
3.3.2 Coherència en el subespai	16
3.4 Transformacions del sistema de referència	16
3.4.1 Traslació	17
3.4.2 Rotació	18
3.5 Programació del procediment GREF1	18
3.5.1 Aportacions al procediment GREF1 (treball futur)	19
3.6 Aplicació del procediment GREF1	20
3.6.1 Resultats del procediment GREF1	20
3.6.2 Interpretació dels resultats	23
3.7 Una altra aplicació del procediment GREF1	24
4. Un altre exemple complet	27
4.1 Presentació de les dades	27
4.2 Selecció d'individus i de variables	27
4.3 ACP enquesta de referència	28
4.4 Selecció del subespai de representació	29
4.5 Inserció de dades	30
4.6 Conclusions	32
5. Conclusions	33

Annex1. Programa MIVAR complet	34
Annex2. Programa GREF1 complet	39
Annex3. Resultats de la inserció per a 3.7	44
Annex3.1 Subconjunt mínim de variables comunes (MIVAR)	44
Annex3.2 Estadístiques descriptives de les variables contínues comunes	44
Annex3.3 Test de correlacions	44
Annex3.4 Coordenades de les variables contínues puntuals	45
Annex3.5 Coordenades dels individus puntuals	45
Annex3.6 Coordenades de l'individu "centre de gravetat"	45
Annex3.7 Coordenades de la variable categòrica	45
Annex4. Resultats de la inserció per a 4	46
Annex4.1 Subconjunt mínim de variables comunes (MIVAR)	46
Annex4.2 Estadístiques descriptives de les variables contínues comunes	47
Annex4.3 Correlacions entre els eixos	47
Annex4.4 Coordenades de les variables contínues puntuals	47
Annex4.5 Coordenades de l'individu "centre de gravetat"	48
Annex4.6 Coordenades de les variables categòriques	48
Bibliografia	49

1. Introducció

1.1 Objectius

El següent projecte va néixer amb la necessitat d'actualitzar uns procediments dissenyats per ser executats en el paquet estadístic SPAD en una versió anterior a la 4.0, orientada per a entorns DOS. Les versions posteriors a la 3.5 (4.0, 4.5 i 5.0) funcionen exclusivament en entorns de Windows de 32 bits (95, 98, NT...) i les rutines s'han de convertir per tal que no entrin en conflicte amb el sistema operatiu. Els procediments en qüestió són dos d'un conjunt de rutines dissenyades per Roser Rius com a complement de la seva Tesi Doctoral (Inserció de dades d'enquesta) que implementen eines i processos per a realitzar una inserció de dades a partir d'una anàlisi factorial. La primera, anomenada MIVAR (detallada en el capítol 2), s'utilitza com a pas previ a la inserció, mentre que la segona, anomenada GREF1 (detallada en el capítol 3) és la que realitza els càlculs necessaris per a la inserció.

El projecte consisteix, a part de la feina de programació pròpiament dita, en presentar l'eina "inserció de dades d'enquesta" i el seu context, les tècniques d'anàlisi factorials i es detallen els dos procediments que defineixen l'eix central del treball, MIVAR i GREF1, dels quals s'explica la filosofia, la tècnica i el disseny algorítmic proposat per portar-los a terme, així com l'aplicació pràctica sobre dades reals.

Els procediments desenvolupats per a Spad estan dissenyats en llenguatge Fortran, per la qual cosa han estat imprescindibles, a part de l'assignatura d'Anàlisi Multivariant de Dades (1r quadrimestre del 3r curs) on s'expliquen les tècniques i la interpretació de resultats d'una anàlisi factorials i l'ús del paquet estadístic SPAD, les assignatures Fonaments d'Informàtica i Programació (2n quadrimestre del primer curs i 1r del 2n respectivament).

1.2 Les Anàlisis Factorials Descriptives

Les anàlisis factorials descriptives (AFD) són eines estadístiques desenvolupades durant el segle XX que serveixen per resumir grans volums de variables interrelacionades en altres creades a partir de la combinació lineal de les originals que en conserven gran part de la variabilitat, per representar grans conjunts de dades en espais bidimensionals, conservant la màxima informació possible.

L'antecedent més immediat de les anàlisis factorials són les tècniques de regressió, iniciades per Galton i continuades, entre d'altres, per Pearson. Se sol atribuir l'origen de les AFD a C. Spearman, en el seu treball sobre la mesura de la intel·ligència (*General Intelligence, Objectively Determined and Measured* (1904)), i de les anàlisis en components principals (ACP) a H. Hotelling (*Analysis of a complex of statistical variables into principal components* (1933)). Altres pioners en les tècniques d'AFD van ser L.L. Thurstone (*Factorial Studies of Intelligence* (1941)) i Kaiser.

Tot i que molts autors consideren l'Anàlisi en Components Principals (ACP) una eina independent a les anàlisis factorials, en el present informe es tractarà com una branca d'aquesta, en concret, l'ACP es considera una AFD aplicada sobre variables contínues, de la mateixa manera que una anàlisi en correspondències és considerada una AFD aplicada a variables categòriques.

Històricament s'han diferenciat dues filosofies al voltant de les AFD que entren en conflicte, no tant en la metodologia algebraica sinó en l'anàlisi i la interpretació de resultats que ofereixen aquestes eines. Per una banda es situa el cos científic britànic, que en el cas concret de les ACP considera la matriu de correlacions entre variables i la matriu de correlacions (o saturacions) entre factors i variables originals i calcula les "comunalitats" (proporció de variabilitat explicada pels factors comuns en una variable) i les compara amb les "comunalitats" originals (en una ACP estimades per 1) per mesurar el grau de bondat de la representativitat dels eixos. L'altre enfocament, i el seguit en aquest treball ja que és l'implementat en SPAD, és el que proposen els estadístics francesos, consistent a reproduir les dades a partir d'una descomposició en valors propis (que representen la variància que suporta l'eix que representa) i vectors propis (que defineixen l'espai on projectar els individus i les variables) de la matriu originals, obtenint unes noves i "poques" variables per als individus que conserven gran part de la informació.

1.3 Procediment per una ACP

Sigui X la matriu ($n \times p$) amb n individus actius i p variables contínues actives, M la matriu de pesos de les variables (o mètrica en l'espai dels individus (R^p , M)) i N la matriu de pesos dels individus (o mètrica en l'espai de les variables (R^n , N)),

$$X = \begin{pmatrix} x_{11} & \dots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{np} \end{pmatrix} \longrightarrow \text{Descomposició en valors singulars}$$

$M^{1/2} X^t N X M^{1/2}$		$N^{1/2} X M X^t N^{1/2}$		(matrius diagonals)
λ_1	u_1	λ_1	v_1	
λ_2	u_2	λ_2	v_2	
\dots		\dots		
λ_r	u_r	λ_r	v_r	$r = \text{rang}(X)$

$\Lambda = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_r \end{pmatrix}$

$U = \begin{pmatrix} u_1 \\ \vdots \\ u_r \end{pmatrix}$

$V = \begin{pmatrix} v_1 \\ \vdots \\ v_r \end{pmatrix}$

La projecció de les files de la matriu, o dels individus, ψ en les noves variables o factors serà:

$$\psi = X M^{1/2} U \text{ o, a partir de } \phi \text{ i aplicant les relacions de transició:}$$

$$\psi = X M \phi \Lambda^{-1/2}$$

I la projecció de les columnes, és a dir, de les variables ϕ :

$$\phi = X^t N^{1/2} V \text{ o, a partir de } \psi \text{ i aplicant les relacions de transició:}$$

$$\phi = X^t N \psi \Lambda^{-1/2}$$

$$\psi = \begin{pmatrix} \Psi_1 & \Psi_2 & \dots & \Psi_r \\ \Psi_{11} & \Psi_{12} & \dots & \Psi_{1r} \\ \vdots & \vdots & \ddots & \vdots \\ \Psi_{nr} & \Psi_{n2} & \dots & \Psi_{nr} \end{pmatrix} \quad \phi = \begin{pmatrix} \Phi_1 & \Phi_2 & \dots & \Phi_r \\ \Phi_{11} & \Phi_{12} & \dots & \Phi_{1r} \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{pr} & \Phi_{p2} & \dots & \Phi_{pr} \end{pmatrix}$$

Es defineixen dues formalitzacions diferents per a una ACP, amb interpretacions estadístiques particulars, l'ACP normalitzada, consistent a analitzar la matriu X havent-la centrat i reduït ($a_{ij} = (a_{ij} - m_j) / s_j$, sent m_j la mitjana de la variable j i s_j la seva desviació tipus) i l'ACP no normalitzada, només havent centrat la matriu X ($a_{ij} = a_{ij} - m_j$). En una anàlisi normalitzada s'estarà diagonalitzant la matriu de correlacions, sense tenir en compte la dimensió de les variables, interpretant-se les projeccions com a correlacions, mentre que en la no normalitzada es diagonalitzarà la matriu de covariàncies.

1.4 Resultats d'una ACP en SPAD

L'SPAD (Système Pour l'Analyse des Données) és un paquet estadístic especialitzat en anàlisis factorials. Està editat per DECISIA, societat molt relacionada amb la comunitat universitària francesa i treballa sota el punt de vista francès de les AFD. Per una ACP, concretament, els resultats que ofereix el programa un cop definides les variables actives contínues per a l'anàlisi, les variables il·lustratives (contínues i il·lustratives), els individus actius/il·lustratius, la mètrica dels espais i el tipus d'anàlisi (normalitzada o no normalitzada) són els següents:

- Informació sobre les variables que intervenen en l'anàlisi i estadística descriptiva.
- Matriu de covariàncies o correlacions entre les variables actives.
- Valor, inèrcia i histograma dels k primers vectors propis.
- Mesura de les diferències entre valors propis per detectar salts significatius.

· Coordenades de les variables sobre els 5 primers eixos factorials. Gràficament es representen aquestes coordenades com a fletxes des de l'origen, si l'anàlisi és normalitzada dins del cercle unitat ja que la correlació variable/factor més gran possible és 1. La direcció de la fletxa indica creixement de la variable, i els individus amb valors alts d'aquesta se situaran prop de la coordenada de la variable. Una manera de mesurar les relacions entre dues variables està en veure l'angle que dibuixen les seves fletxes en els plans factorials. Un angle petit representarà que la relació es forta i positiva. Un angle pràcticament recte representarà la independència de variables i un angle pla indica que la relació és forta però negativa.

- Coordenades de les variables contínues il·lustratives i de les variables categòriques. Per aquestes últimes es calcula la coordenada sobre cada factor i el valor-test (valor en una distribució normal estàndard) corresponent que contrasta si la distància a l'origen és significativa o no. En cas de ser-ho, s'ha de relacionar cada categoria amb les variables contínues que tenen el mateix signe en aquell eix. Dins del pla factorial, es dibuixen com a punts aïllats i si una variable només té dues categories sempre se situen de forma simètrica.

- Coordenades dels individus (opcional). De la mateixa manera que per a les variables, es calcula les coordenades sobre els eixos dels individus i es poden representar en plans factorials.

- La simplicitat que genera una AFD queda palesa a l'hora d'analitzar gràficament els plans factorials, sobretot quan el nombre d'eixos significatius són pocs, doncs es resumeixen gran quantitat de relacions entre variables en un sol pla. Spad permet il·lustrar diferents tipus de gràfics sobre els eixos factorials i inclús multi-gràfics per a comparar plans originaris d'una AFD dos a dos.

1.5 Inserció de dades d'enquesta

La necessitat de l'eina estadística anomenada “inserció de dades” prové de la voluntat de voler relacionar informació originària de dos estudis independents. Anteriorment s'havia atacat aquesta qüestió mitjançant eines clàssiques com la regressió o la reponderació de les mostres i des del punt de vista no probabilístic o d'anàlisis factorials, sempre havia existit el problema d'imputació de dades.

Existeixen gran volum de conjunts de dades provinents de naturalesa molt diferent però amb un conjunt de blocs d'informació comuns que formen uns lligams capaços d'extendre's entre les informacions no comunes. En concret, considerarem normalment una enquesta de referència, un estudi exhaustiu, anterior a una enquesta puntual, amb informació més particular i ens interessarà poder establir connexions entre aquestes dos conjunts de dades a partir de la informació comuna (aquella que ambdós estudis recullen). S'anomena espai comú (R^p) el conjunt de variables comunes en les dues enquestes. En aquest espai es poden representar els n_0 individus corresponents a l'enquesta de referència i els n_1 corresponents a l'enquesta puntual (veure figura 1.1) i a través d'aquests, visualitzar les variables comunes i no comunes (inserir). Per a determinar quines són les variables que convé tornar a recollir en el segon estudi perquè aquest contingui la informació que permeti formar lligams fiables existeix el procediment MIVAR (veure capítol 2).

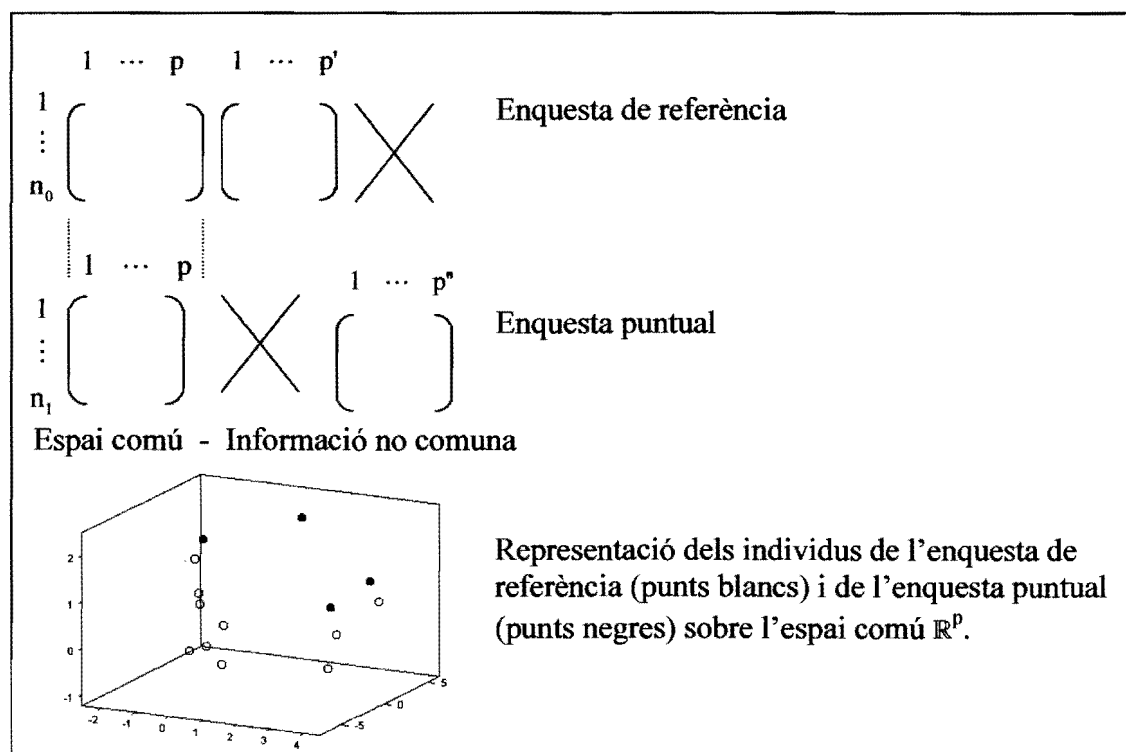


Figura 1.1. Esquema de l'estructura de dades que plantegen el problema d'insertir

Existeixen dues maneres de definir la inserció: la inserció en referència (veure figura 1.2), tècnica sobre la qual es tractarà en aquest informe, utilitzada quan ens interessa conservar el marc de referència (eixos factorials determinats per una AFD sobre l'enquesta de referència) que defineix l'enquesta primera, on hi voldrem representar les dades de l'enquesta puntual (sempre es considera que l'enquesta de referència és un estudi més exhaustiu que la puntual, i per tant, més fiable) i la inserció conjunta, utilitzada quan el marc de referència no presenta les característiques adequades per dur a terme l'anterior tipus, per exemple, perquè en l'enquesta puntual és impossible reproduir el mateix subespai de representació.

Un cop escollit el subespai de representació a través d'una AFD de les dades de l'enquesta de referència (per a aquest treball es limita al cas que el subespai és generat per una ACP i per tant, l'espai comú és format per variables contínues), la tècnica d'insertió d'individus es basa en la projecció en suplementari i la inserció de variables en la projecció doblement suplementària, proposades per Pagès, Bonnefous i Iliakopoulos adonant-se de la repetició d'informació entre les enquestes que es realitzaven a França.

La projecció en suplementari consisteix en l'aplicació de les fórmules anteriorment vistes en l'ACP per a traslladar els individus o les variables a l'espai que formen els eixos factorials per a individus que no formen part del càlcul d'aquest espai, i la doblement suplementària en l'aplicació de les relacions de transició, que permeten calcular la projecció ϕ com a funció de ψ . El procediment GREF1 (veure capítol 3) s'encarrega d'aquestes operacions.

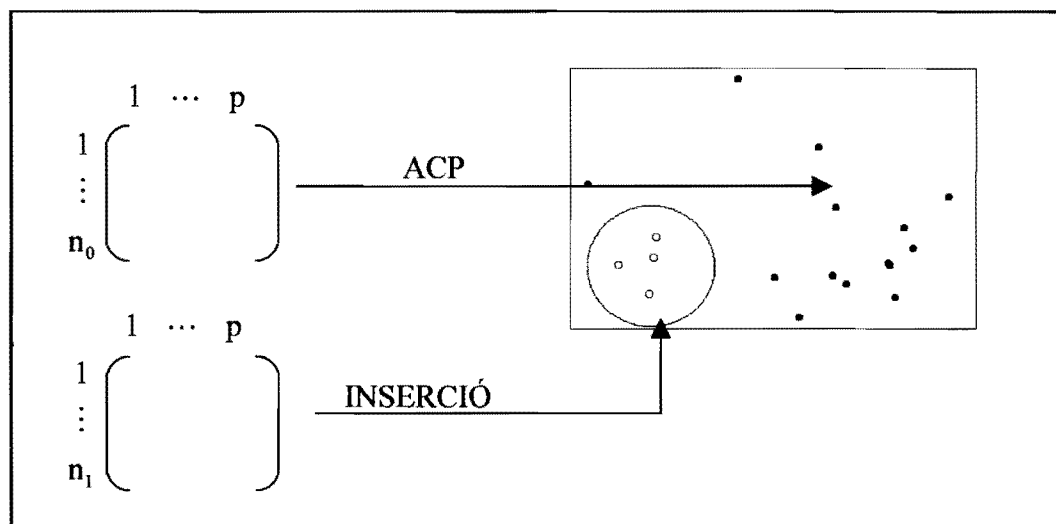


Figura 1.2. Esquema de funcionament de la inserció en referència per a individus

2. Procediment MIVAR

2.1 Plantejament

El procediment MIVAR (conjunt Mínim de VARIables) serveix per a determinar el subconjunt òptim de variables de l'enquesta de referència per a reproduir un marc en el qual poder fer la inserció de dades de l'enquesta puntual. Les variables escollides seran les que s'hauran de recollir en la segona enquesta per tal que la inserció es dugui a terme en un marc adequat. Com que l'enquesta puntual és, per definició, més petita que la de referència, no pot ser plantejat tornar a recollir totes les variables una altra vegada. Aquest procés es tradueix a buscar el subconjunt de variables que millor conserva l'estructura inicial de l'ACP segons un criteri establert. El procediment dissenyat elimina una variable en cada passada de l'algoritme (no garantint, per tant, l'òptim global) per minimitzar el cost computacional, seleccionant per eliminar aquella que menys utilitat té en la formació d'un marc per a l'inserció. Partint de p variables (les considerades actives en l'anàlisi factorial sobre l'enquesta de referència) es calculen k eixos factorials, definits per l'usuari, per les p combinacions possibles de $p-1$ variables i les correlacions entre els eixos originals i els reconstruïts havent eliminat una variable. Aquella combinació que maximitza la suma de les correlacions al quadrat ponderada pel valor propi de l'eix factorial reconstruït serà l'escollida per a formar el nou marc de referència. Es repeteix el procediment fins a arribar al subconjunt de cardinalitat mínima, també definida per l'usuari. El procediment també reconstrueix el marc complet, afegint pas a pas la variable que maximitza el criteri per considerar altres òptims locals que es poden generar amb les variables que disposem.

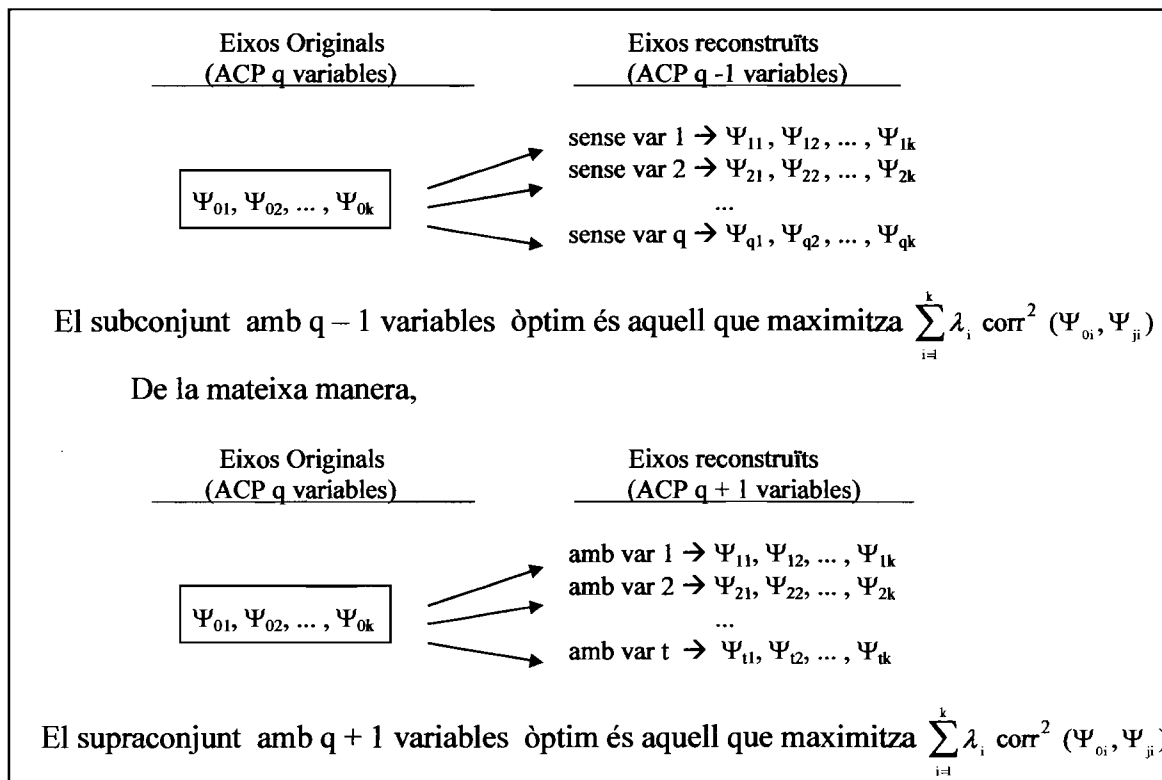


Figura 2.1. Funcionament teòric del procediment MIVAR

2.2 Disseny del procediment

La solució proposada per a implementar el procediment MIVAR es basa en una cerca "greedy", que en cada passada del bucle principal elimina (o inclou) una de les variables. Els paràmetres d'entrada que necessita el programa, a part de l'adreça del fitxer NDIC i NGUS corresponents a la descripció de variables i individus de les dades i les coordenades sobre els eixos factorials originaris d'una ACP amb totes les variables actives, el nombre mínim de variables a què arribar (NMVAR) i el nombre d'eixos sobre els quals es calcularà el criteri (NAXR). Si no es pot generar tants eixos com els especificats perquè es demanen més eixos que de variables que disposem, la cardinalitat mínima serà el nombre d'eixos especificat.

A continuació s'explicita el disseny bàsic en pseudocodi del programa:

```
algoritme mivar és
obrir_fitxer (NDIC)    // fitxer diccionari de dades
obrir_fitxer (NGUS)    // fitxer de coordenades en eixos factorials AFD de vars i individus
llegir_paràmetres (NAXR, NMVAR)
si (NAXR > NMVAR) llavors NMVAR = NAXR fsi
mentre (no NMVAR) fer
  per (cadascuna de les variables actives) fer
    escriure_fitxer (NDIC, NDON)    // excloent una variable
    calcular_estadístiques
    calcular_correlacions
    diagonalitzar_matriu
    calcular_distàncies
    calcular_correlacions_entre_eixos    // corr ( $\Psi_{0,i}$ ,  $\Psi_{j,i}$ )
    calcular_criteri                    //  $\sum_{i=1}^k \lambda_i \text{corr}(\Psi_{0,i}, \Psi_{j,i})$ 
  fper
  escriure_resultats                    // variable a eliminar i criteri conservat
fmentre
mentre (no nombre total de variables) fer
  per (cadascuna de les variables actives) fer
    escriure_fitxer (NDIC, NDON)    // incloent una variable
    calcular_estadístiques
    calcular_correlacions
    diagonalitzar_matriu
    calcular_distàncies
    calcular_correlacions_entre_eixos    // corr ( $\Psi_{0,i}$ ,  $\Psi_{j,i}$ )
    calcular_criteri                    //  $\sum_{i=1}^k \lambda_i \text{corr}(\Psi_{0,i}, \Psi_{j,i})$ 
  fper
  fmentre
  escriure_resultats                    // variable a afegir i criteri millorat
falgoritme
```

En l'Annex es detalla el programa complet i les subrutines necessàries en codi Fortran preparat per a la compilació i execució per la versió de desenvolupament d' SPAD.

2.2.1 Aportacions al procediment MIVAR

Arran de la conversió del procediment, ha sorgit també l'oportunitat de millorar alguns aspectes del procediment MIVAR, sobretot referents a la sortida en pantalla. En la primera versió, els resultats que oferia la rutina consistien en imprimir, per a cada passada del bucle per a eliminar una variable, el criteri calculat per a cadascuna de les variables possibles per a eliminar i finalment, l'etiqueta de la variable que maximitza el criteri i per tant, s'eliminarà. Aquesta solució, a part de generar molta informació i text prescindible, és molt completa localment però incòmoda d'analitzar globalment, i s'ha decidit modificar-la.

En l'actualitat i com es podrà veure en les aplicacions més endavant, el procediment abans de començar el bucle d'eliminació ofereix una taula amb les variables contínues actives en l'ACP i que es disposen a eliminar i el criteri i la llista de valors propis inicials. Per a cada passada del bucle, escriu per pantalla la variable que s'elimina i el criteri, tant numèricament com gràfica (en relació amb el criteri inicial). A l'hora de remuntar els resultats són els mateixos, fins a arribar al criteri inicial una altra vegada. D'aquesta manera, tot i perdre informació, és més senzill seguir quin procés a seguit el programa per a arribar al subconjunt mínim i mitjançant l'histograma de criteris es pot veure si hi ha un salt significatiu que representaria una gran pèrdua ocasionada per l'eliminació de la variable en qüestió.

2.3 Aplicació procediment MIVAR

Per a posar a prova el procediment s'utilitzaran unes dades subministrades pel paquet estadístic SPAD referents a una enquesta realitzada a 315 persones amb 12 variables contínues: edat, nombre d'individus a la llar, nombre de fills, ajudes familiars, ingressos personals i 7 opinions sobre la importància que es dona a diversos conceptes. Seran les 7 variables referents a opinió les que s'utilitzaran com a actives en l'ACP per centrar l'anàlisi sobre la opinió de la mostra. A més a més, l'enquesta recull 38 variables categòriques que de moment no jugaran cap paper. Podria sorgir la necessitat de preguntar a un altre col·lectiu de gent independent a l'anterior sobre la importància que es dona als mateixos o altres temes, a part d'altres qüestions. Per a poder relacionar ambdós estudis cal una recollida de dades en el segon que se sobreposi amb el primer, però potser no totes 7 variables són imprescindibles per a crear un subespai de referència bo on inserir els individus i les variables i tot el que suposi una reducció de la despesa mereix atenció especial.

Dividirem el conjunt de dades en dos, un referent a l'anomenada en termes d'inserció, enquesta de referència, i l'altre que farà el paper d'enquesta puntual. L'enquesta de referència, per definició més exhaustiva que la puntual estarà formada pels 265 primers individus i inclourà les 7 variables d'opinió, edat, nombres d'individus a la llar, i ingressos de l'enquestat i 2 variables categòriques de les 38 (per simplificar l'exemple): sexe i disposició d'un piano. L'enquesta puntual, suposadament posterior, recollirà per a la resta d'individus les variables d'opinió que ens determini el procediment MIVAR, la resta de variables contínues que no hem agafat per a la de referència i dos variables categòriques més: possessió de segona residència i haver patit un atac de nervis recentment.

2.3.1 Anàlisi en components principals de l'enquesta de referència

Es disposa a analitzar de manera usual l'enquesta de referència, per determinar quants eixos són necessaris i interpretar-los.

ACP normalitzat

Individus actius: Els 265 primers

Variables actives: Opinió sobre fills (41), feina (42), oci (43), amistats (44), pares, germans i sogres (45), religió (46) i política (47)

Variables il·lustratives: edat (4), nombre d'individus (26) a la llar i ingressos de l'enquestat (51) (contínues) i sexe (3) i possessió d'un piano (16) (categòriques)

En l'ANNEX es mostra els resultats oferts per SPAD, d'on destaquen les següents conclusions:

- Les variables contínues més correlacionades són l'opinió sobre fills i sobre pares, germans i sogres ($\text{corr}(41,45) = 0,24$), opinió sobre fills i treball ($\text{corr}(41,42) = 0,19$), entre feina i oci ($\text{corr}(42, 43) = 0,20$), amistat i oci ($\text{corr}(43, 44) = 0,32$), amistat i pares, germans i sogres ($\text{corr}(44, 45) = 0,21$) i entre religió i política ($\text{corr}(46,47) = 0,31$).

- L'histograma de valors propis mostra un decreixement exponencial que es dispara sobretot entre el quart i el cinquè. Els quatre primers valors propis representen un 72,75 % de la inèrcia total.

- El primer eix factorial és un factor de creixement, ja que totes les variables actives hi tenen la coordenada en positiu i separa els individus que a tot li donen molta importància dels qui en donen menys també a totes les variables. Les més correlacionades a aquest factor són la importància concedida a l'oci, l'amistat i als pares, germans i sogres (correlacions de l'ordre de 0,60). El segon eix factorial (fortament lligat a la variable corresponent a la religió) separa els individus que donen molt valor a l'amistat a la religió i a la política en contraposició de qui ho fan per la resta de variables. El tercer eix contraposa els qui donen importància a l'oci i a l'amistat dels qui ho fan a la resta de temes i el quart eix factorial als qui importa la feina, l'oci, la religió i la política de als qui només importa la família i els amics.

- De les variables contínues il·lustratives s'observa que el factor de creixement funciona igual per a totes elles, que la gent gran està fortament lligada al segon eix (donen més importància a la religió, la política i els amics) mentre que com més gent hi ha a casa, més importància es donen a temes familiars i de feina i tendeixen a donar poca importància a l'oci i a l'amistat.

- Per què fa a les variables categòriques veiem que el sexe només és significatiu en el tercer eix factorial (els homes donen més importància a l'amistat i a l'oci que no a la resta de temes i les dones al contrari) i la possessió del piano no té cap relació amb l'opinió dels enquestats.

2.3.2 Resultats del procediment MIVAR

Havent arribat a la conclusió que els 4 primers eixos són aquells que conserven la major part de la variabilitat de les dades, obtenim el primer paràmetre pel procediment MIVAR, doncs utilitzarem aquests 4 eixos per a calcular el criteri que ens ha de proporcionar quina variable s'elimina en cada pas de l'algoritme. El nombre mínim de variables que volem conservar ja depèn d'altres factors, com pot ser la decisió del responsable de l'estudi. No obstant, si es vol conservar 4 eixos, és imprescindible conservar alhora 4 variables per a generar-los i per tant, es considera que el nombre de variables mínim sigui 4 igualment. Els resultats que ofereix el procediment són els següents:

```

PROCEDURE DUMMY - 1 -
RECONSTRUCTION DES      4 PREMIERS AXES
CRITERE INITIAUX AVEC    7 VARIABLES:      5.093
(somme des correlations carrees des premiers axes ponderees par les valeurs propres)
VALEURS PROPRES DES      4 PREMIERS AXES:
1.6290545
1.3419726
1.1421208
0.9795322
LISTE DES 7 VARIABLES INITIAUX:
1:Fami Famille, enfants : importance accordée
2:Trav Travail, profession : importance accordée
3:Lois Temps libre, détente : importance accordée
4:Amis Amis, connaissances : importance accordée
5:Part Parents, frères, soeurs ... : importance accordée
6:Reli Religion : importance accordée
7:Poli Politique, vie politique : importance accordée
ELIMINATION DES VARIABLES PAS A PAS JUSQUA 4
Step Var. eliminee Critere
7 -> 6 VARS. 5 3.945 *****
6 -> 5 VARS. 1 2.436 *****
5 -> 4 VARS. 6 2.076 *****
INCORPORATION DES VARIABLES PAS A PAS JUSQUA 7
Step Var. ajoutee Critere
4 -> 5 VARS. 6 2.436 *****
5 -> 6 VARS. 1 3.945 *****
6 -> 7 VARS. 5 5.093 *****

```

Figura 2.2. Resultats impresos pel procediment MIVAR amb les dades de l'exemple

Aquests resultats s'interpreten com que el subconjunt de 4 variables que millor conserva la identitat de l'ACP original amb totes les variables és el format per les variables enumerades amb 2, 3, 4 i 7 corresponents a la importància que es dona a la feina, al temps lliure, a les amistats i a la política. La variable menys útil per a formar part del subespai de representació i la primera en ser eliminada és la importància que es dona als pares, germans i sogres (5), amb una disminució del criteri respecte l'original de (23%), seguida de la importància que es dona als fills (6) (amb una disminució força considerable ja del criteri, en un 52%) i a la religió (1) (disminució del 59%). La primera variable a eliminar és l'opinió sobre els pares, germans i sogres per l'alta correlació que existeix entre aquesta variable i la d'opinió sobre fills. Conservant una de les dues variables ja se'n té prou per recollir la informació que ens interessa. També destaca de l'anàlisi dels resultats que amb més de la meitat de les variables no conservem el 50% del criteri inicial i per tant sembla evident que amb quatre variables no en tenim prou per formar un bon subespai de representació.

A l'hora de remuntar variables fins a formar el marc original, la inclusió pas a pas coincideix amb l'exclusió, introduint primer la variable 6, després la 1 i per últim la 5. Arran dels resultats suscita el debat de quin és el nombre de variables adequat, aquell que no suposi una gran pèrdua i, com s'ha comprovat, eliminant la primera variable ja es perd un 23% del criteri, per la qual cosa, sembla raonable conservar totes les variables en l'enquesta puntual.

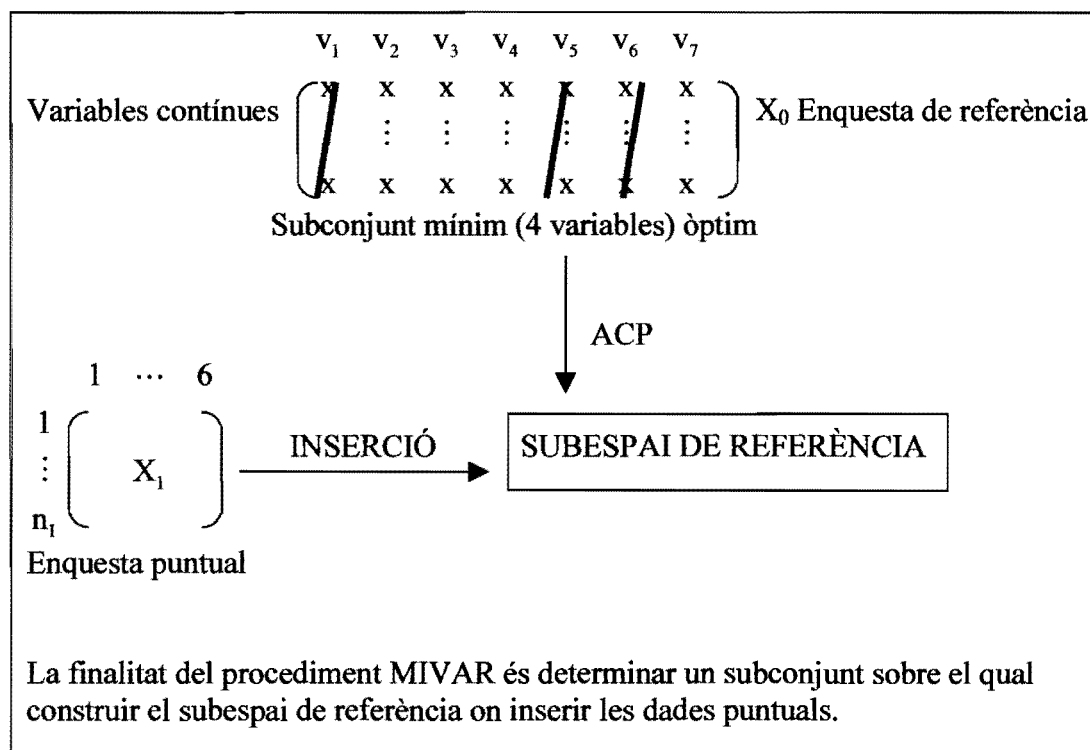


Figura 2.3. Implicacions del procediment MIVAR si el subconjunt mínim és representatiu. En l'exemple no es dona el cas i es conformarà el subespai de representació amb totes 7 variables

3. El procediment GREF1

3.1 Plantejament

Mitjançant el procediment GREF1 (GREFFer, inserir en francès) s'implementa la inserció en referència pròpiament dita de les dades de l'enquesta puntual al subespai de representació determinat a partir de l'espai comú a una primera enquesta, tal i com s'ha definit en l'apartat 1.3. Ens ocuparem del cas en què l'AFD precedent és una ACP i sense aplicar translació ni rotació, tècniques per a garantir la coherència de les dades a inserir descrites més endavant.

La inserció en referència es basa en la projecció en suplementari dels individus i la projecció doblement suplementària de les variables. Partint d'una AFD normal de l'enquesta de referència per a determinar el subespai de representació on projectar (es recorda que és en l'elecció d'aquest subespai on entra el joc el procediment MIVAR, per eliminar variables poc útils), obtenim els vectors propis $u_{0\alpha}$, amb els corresponents valors propis λ_α , on es projectaran els individus i les variables de l'enquesta de referència de forma usual. La projecció en suplementari consisteix a projectar de la mateixa manera els individus de l'enquesta puntual. Sigui X_1 la matriu de dades formada pels individus de l'enquesta puntual, la projecció es realitza segons la fórmula següent:

$\Psi_{1\alpha}^{\text{sup}} = X_1 M_0^{1/2} u_{0\alpha}$, on M_0 és la matriu de pesos de les variables comuns definida per a l'enquesta de referència i $u_{0\alpha}$ és l' α -èssim vector propi sorgit de l'ACP de l'enquesta de referència

Fent ús de les definides relacions de transició, es projecta doblement en suplementari per a les variables:

$\Phi_{1\alpha}^{\text{ds}} = \frac{1}{\sqrt{\lambda_\alpha}} X_1^t N_1 \Psi_{1\alpha}^{\text{sup}}$, on N_1 és la matriu de pesos dels individus definida per a l'enquesta puntual i λ_α l' α -èssim valor propi de l'ACP de l'enquesta de referència

Aquesta projecció s'anomena doblement suplementària perquè per a representar les variables en el marc de l'enquesta de referència s'utilitza la projecció dels individus, calculada en el seu moment en suplementari. Fixi's que per a projectar els individus s'utilitza la matriu de pesos per a les variable de l'enquesta de referència mentre que a projectar les variables de l'enquesta puntual, la matriu de pesos dels individus és la de l'enquesta puntual, també.

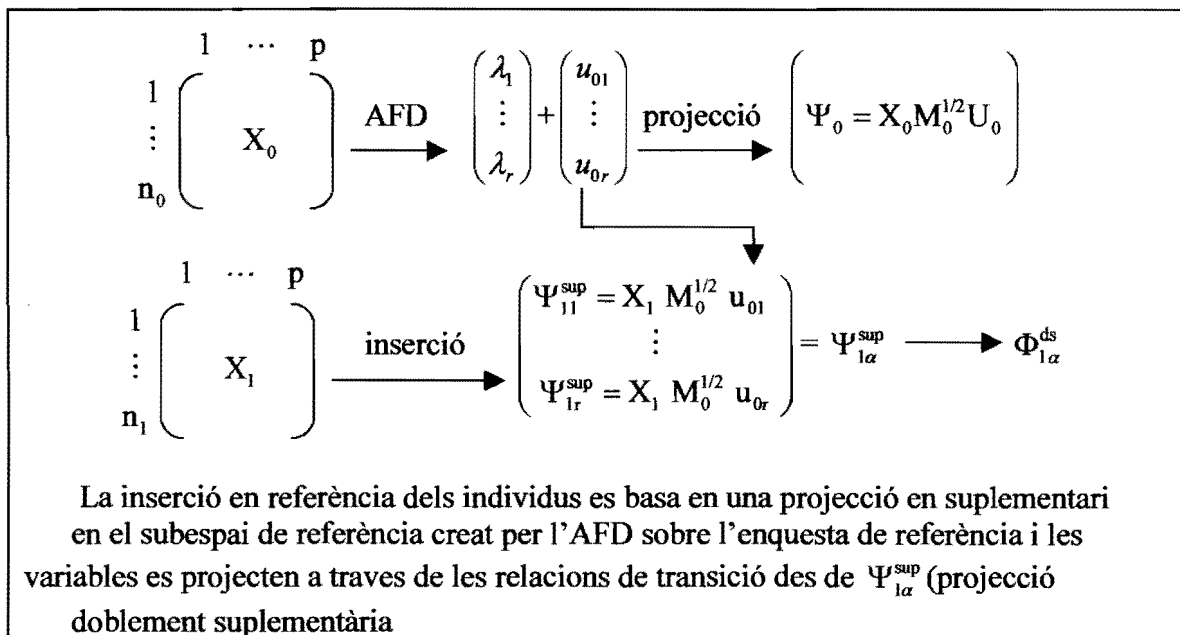


Figura 3.1. Esquema del funcionament teòric de la inserció

3.2 Inserció a partir d'una ACP

Per al cas que ens ocupa, quan l'AFD aplicada és una ACP, siguin X_0 la matriu ($n_0 \times p$) de dades provinents de l'enquesta de referència, amb p variables contínues centrades amb les mitjanes dels n_0 individus (si l'anàlisi fos normalitzat, es redueix la matriu amb les desviacions tipus), C_1 la matriu ($n_1 \times p$) de dades originàries de l'enquesta puntual amb p variables contínues centrades amb la mitjana dels individus d' X_0 (per a situar el núvol de dades a inserir amb el núvol de dades de referència) per a n_1 individus independents als anteriors i X_1 la matriu ($n_1 \times p$) de dades puntuals centrada i reduïda amb les mitjanes i les desviacions tipus de les dades de referència ($X_1 = C_1 M_0^{1/2}$ on M_0 és la matriu diagonal que conté les inverses de les variàncies de les p variables de l'enquesta de referència), la projecció en suplementari dels n_1 individus a inserir es calcula mitjançant:

$$\Psi_{1\alpha}^{sup} = X_1 u_{0\alpha} = C_1 M_0^{1/2} u_{0\alpha} \text{ si l'ACP és normalitzada}$$

o

$$\Psi_{1\alpha}^{sup} = C_1 u_{0\alpha} \text{ si l'ACP és no normalitzada}$$

Per a la projecció de les variables contínues (actives i il·lustratives) s'utilitza la projecció doblement suplementària però adaptada per a que les coordenades coincideixin amb la correlació entre les variables i els eixos factorials. Sigui Y_1 la matriu de dades a inserir amb les variables centrades i reduïdes amb les seves pròpies mitjanes i desviacions tipus i W_1 les dades a inserir amb les variables centrades però no reduïdes, i $g_{1\alpha}^{sup}$ i $\text{stdev}(\Psi_{1\alpha}^{sup})$ la mitjana i la desviació tipus de $\Psi_{1\alpha}^{sup}$ respectivament,

$$\begin{aligned}
\Phi_{1\alpha}^{ds} &= \frac{1}{\text{stdev}(\Psi_{1\alpha}^{sup})} \sum_{i=1}^{n_1} \frac{1}{n_1} y_{1ij} (\Psi_{1ai}^{sup} - g_{1\alpha}^{sup}) \\
&= \frac{1}{\text{stdev}(\Psi_{1\alpha}^{sup})} \frac{1}{\text{stdev}(w_1^j)} \frac{1}{n_1} \sum_{i=1}^{n_1} (w_{1ij} - 0) (\Psi_{1ai}^{sup} - g_{1\alpha}^{sup}) \\
&= \text{corr}(w_1^j, \Psi_{1\alpha}^{sup}) \\
&= \text{corr}(y_1^j, \Psi_{1\alpha}^{sup})
\end{aligned}$$

Per a projectar les variables categòriques puntuals, sigui S_1 la matriu disjuntiva completa de les modalitats dels n_1 individus, la fórmula de la projecció doblement suplementària correspon a calcular per a cada modalitat j la coordenada com a centre de gravetat ponderat dels individus que hi pertanyen:

$$\Gamma_{1aj}^{ds} = \sum_{i=1}^{n_1} \frac{1}{n_1} s_{1ij} \Psi_{1ai}^{sup} = \sum_{\{i | s_{1ij}=1\}} \frac{1}{n_{1j}} \Psi_{1ai}^{sup}$$

3.3 Condicions de validesa de la inserció

S'està definint la inserció en referència, com la projecció d'uns individus i unes variables comunes i no comunes d'una enquesta puntual a un espai generat per una enquesta de referència, independent a l'anterior. Per portar a terme aquesta operació és necessari que els dos conjunts de dades presentin unes característiques "semblants" que es contrasten mitjançant les dues eines següents.

3.3.1 Coherència global : Test de correlacions

Per comprovar la coherència global de les estructures d'ambdues enquestes es proposa realitzar un test que compari les interrelacions de les variables que generen l'espai comú tant per als individus de referència com els puntuals.

Quan es pretén realitzar una inserció de dades provinent d'una ACP normalitzada, podem comparar els elements de les dues matrius de correlacions corresponents en el subespai comú per a les dues enquestes, concretament, comparant dos a dos els coeficients de correlació entre variables en els dos estudis. En el seu moment es va decidir realitzar la prova per a elements dos a dos i no entre matrius per l'interès que desperta saber on existeixen les diferències entre estudis. Siguin r_0 un element de la matriu de correlacions de les dades de referència i r_1 l'equivalent en les dades puntuals, sota la hipòtesis nul·la $H_0 : \rho_0 = \rho_1$ es compleix que:

$$\frac{z_1 - z_0}{\sqrt{\left(\frac{1}{n_1 - 3} + \frac{1}{n_0 - 3}\right)}} \sim N(0,1) \quad \text{on} \quad z_k = \frac{1}{2} \log \left(\frac{1 + r_k}{1 - r_k} \right)$$

El procediment GREF1 realitza, si així es desitja, el test de correlacions i indica si per cada parella de variables el coeficient de correlació són estadísticament iguals en les dues enquestes.

L'equivalent al test de correlacions per a anàlisi en correspondències és l'anomenat test d'homogeneïtat, consistent en fer el test de comparació de les taules creuades de les variables dues a dues, és a dir, en contrastar la homogeneïtat de les distribucions d'efectius de totes les variables comunes dues a dues.

3.3.2 Coherència en el subespai

Un cop determinat si les variables tenen un comportament dos a dos semblant en les dues enquestes o no, ens pot interessar també esbrinar si el subespai de representació on s'insertaran les dades puntuals té coherència amb el propi espai que generarien elles mateixes. Per a contrastar aquesta característica s'aplica una comparació entre vectors propis generats per l'ACP de l'enquesta de referència i els que generaria una ACP de l'enquesta puntual, calculant-ne les correlacions per als k eixos de representació escollits.

És desitjable que aquestes correlacions estiguin dins d'un interval calculat a partir de l'enquesta de referència mitjançant les correlacions dels vectors propis creats per repetició de l'ACP amb dades pertorbades per a estimar l'estabilitat del subespai de representació. L'eina utilitzada per a estimar l'estabilitat del subespai de referència es basa en el mètode bootstrap i està implementada en el procediment STAPC, el qual no es tracta en aquest projecte.

Si les diferències entre correlacions resultessin molt significatives caldrà transformar les dades puntuals mitjançant una rotació, tècnica que s'explica en el següent apartat.

3.4 Transformacions del sistema de referència

La inserció en referència es basa en la projecció en suplementari dels individus i la projecció doblement suplementària de les variables puntuals en el subespai de representació determinat per l'ACP (AFD en general) de l'enquesta de referència. Pot donar-se el cas que el núvol de punts dels individus puntuals inserits en els plans factorials no tinguin com a centre de gravetat l'origen de coordenades, com tenen els individus de referència i en general els individus actius en una AFD. Alhora, quan falla la coherència en el subespai de representació, s'ha vist que els vectors propis de l'ACP no representin els individus puntuals com es desitjaria que ho fessin. Aquests dos problemes poden solucionar-se senzillament mitjançant transformacions rígides: la translació i la rotació.

3.4.1 Traslació

La translació és una tècnica adient quan el núvol de punts dels individus inserits (amb centre de gravetat g_1) no estan centrats en l'origen (o centre de gravetat dels individus de referència g_0) i pels quals els valors-test de les modalitats de l'enquesta de referència tindran un sentit diferent ja que es distribueixen al voltant del centre de gravetat definit pels individus de referència.

Per mesurar la distància entre núvols, es genera un nou individu "centre de gravetat" (dels individus puntuals) pel qual es calcularà la coordenada en els eixos factorials on s'insereix i el valor-test per determinar si és significativament diferent al centre de gravetat dels individus de referència, en general zero. Quan es doni el cas que els dos núvols tenen centres de gravetat estadísticament diferents es pot optar per aplicar una inserció amb translació, que consisteix en centrar el núvol inserit respecte al de referència.

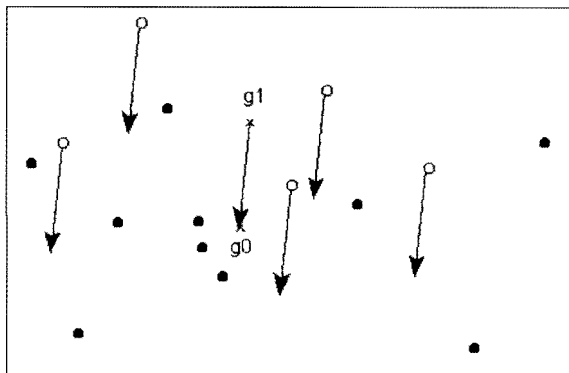


Figura 3.2 Distribució dels individus de referència (punts negres) i puntuals (punts blancs) amb els respectius centres de gravetat (g_0 i g_1) en un pla factorial i l'efecte que hi té una translació

La inserció en referència amb translació (no implementada en la versió de GREF1 dissenyada per al projecte) re-calcula les coordenades dels individus i les modalitats de variables categòriques restant a les que ofereix la projecció en suplementari i doblement suplementària el centre de gravetat dels individus:

$$\Psi_{1_a}^{\text{sup}} \leftarrow \Psi_{1_a}^{\text{sup}} - g_{1_a}^{\text{sup}} \quad \Gamma_{1_a}^{\text{sup}} \leftarrow \Gamma_{1_a}^{\text{sup}} - g_{1_a}^{\text{sup}}$$

La translació no té efectes sobre la coherència global ni la coherència en el subespai, només en les coordenades dels dos elements esmentats, que s'interpretaran tenint en compte que s'ha eliminat el canvi de mitjanes de les variables d'ambdós estudis.

3.4.2 Rotació

La rotació s'aplica per a resoldre un problema de coherència en el subespai, quan el núvol puntual en l'espai comú dona una estructura d'interrelacions de variables que difereix en un angle considerable de l'estructura donada pel núvol de referència. Això indica que el comportament de les variables és semblant (el test de correlacions no donarà cap diferència significativa) però que existeix cert canvi de rotació, que en solucionar, millorarem les característiques de coherència necessàries per a validar la inserció.

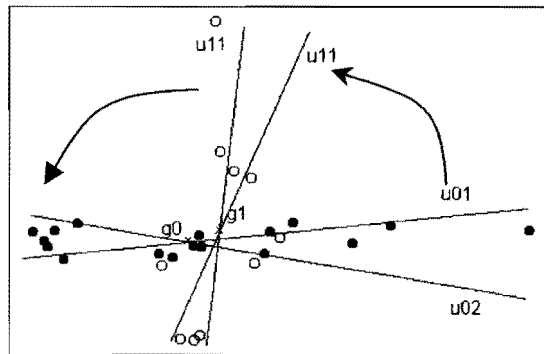


Figura 3.3 Representació del núvol de punts puntual (punts blancs) i de referència (punts negres) amb dos vectors propis corresponents i l'efecte que hi tindria una rotació

Aplicar la rotació no és tan senzill com aplicar una rotació ja que la transformació que es fa de les dades és molt (en alguns cops massa) profunda.

3.5 Programació del procediment GREF1

L'algoritme implementat per a comprovar les condicions de validesa i calcular la inserció té la següent estructura. Els paràmetres d'entrada que requereix són LEDIN (indica si es vol donar les coordenades per als individus), NAXED (nombre pels que es dona les coordenades) i LTEST (que indica si es vol realitzar el test de correlacions):

algoritme gref1 és

```

obrir_fitxer (NDIC) // diccionari de dades de l'enquesta de referència
obrir_fitxer (NDON) // dades de l'enquesta de referència
obrir_fitxer (NGUS) // coordenades sobre eixos factorials dades de l'enquesta de referència
obrir_fitxer (NDIC1) //diccionari de dades de l'enquesta puntual
obrir_fitxer (NDON1) // dades de l'enquesta puntual
llegir_paràmetres (LEDIN, NAXED, LTEST)
fer_estadístiques_descriptives_enquesta_referència
centrar_i_reduir (NDON)
fer_estadístiques_descriptives_enquesta_puntual
centrar_i_reduir (NDON1)
si LTEST=1 llavors
    realitzar_test_correlacions
fsi
calcular_correlacions_entre_eixos
calcular_projecció_per_individus_puntuals
si LEDIN ≠ 0
    mostrar_coordenades_individus
fsi
calcular_i_mostrar_projecció_centre_gravetat_individus
calcular_i_mostrar_projecció_variables_puntuals //continues i categòriques
falgoritme

```

3.5.1 Aportacions al procediment GREF1 (treball futur)

Cal tenir en compte que per a la versió Windows d'SPAD (per icones) les successions d'instruccions (filières) només poden provenir d'una matriu de dades (base) per la qual cosa no es pot accedir a dos fitxers NDIC, NDON si aquests no han estat creats prèviament per un procediment independent a partir d'una sola matriu. Aquest problema és evitable si es treballa des de l'executador de comandes (script), no obstant queda pendent la realització del procés que generi a partir d'una base dos anàlisis independents.

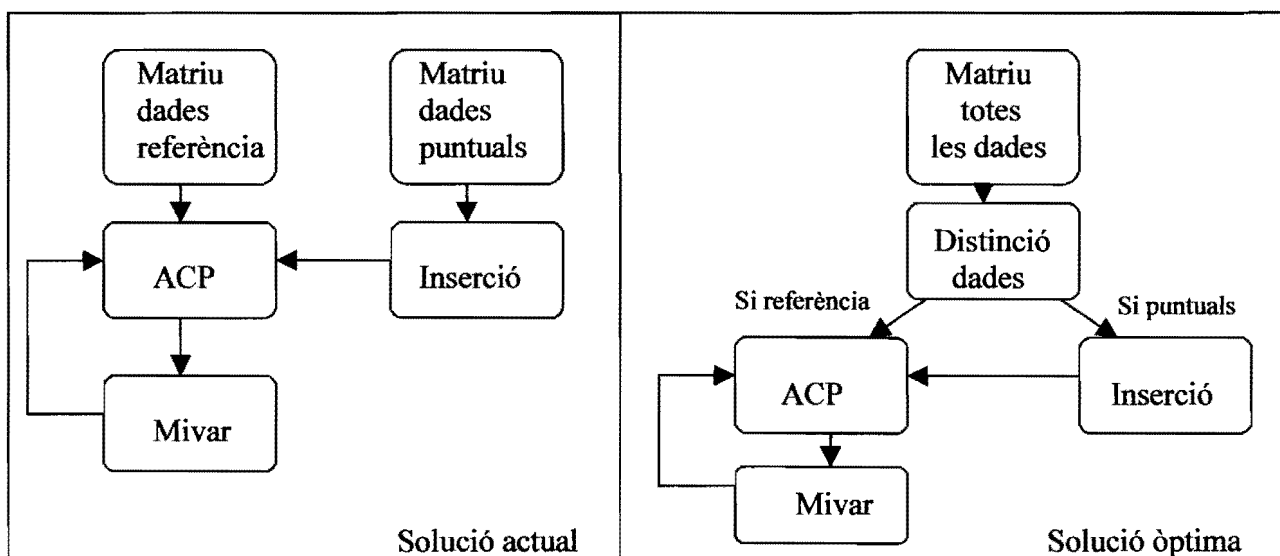


Figura 3.4. Diagrama dels procediments previs a la inserció en la versió actual i en una possible versió futura, optimitzada per a la versió Windows d'Spad

3.6 Aplicació del procediment GREF1

Es procedirà a assajar el procediment GREF1 sobre les dades presentades per a aplicar el procediment MIVAR corresponents a una enquesta sobre 315 individus que s'han dividit en dues enquestes, que fan el paper d'enquesta de referència i enquesta puntual. S'enfocarà el problema havent obtingut per a l'enquesta puntual les 7 variables contínues actives referents a opinió que també inclou l'enquesta de referència ja que mitjançant el procediment MIVAR hem determinat que totes i cadascuna d'elles és imprescindible per a crear un subespai de representació vàlid, deixant com a il·lustratives les variables 'nombre de fills' i 'ajudes familiars' i les variables categòriques 'possessió de segona residència' i 'patir dels nervis'.

3.6.1 Resultats del procediment GREF1

La impressió més completa de resultats que ofereix el programa és:

STATISTIQUES SOMMAIRES DES VARIABLES CONTINUES							
EFFECTIF TOTAL :		265	POIDS TOTAL :		265.00		
NUM .	IDEN - LIBELLE	EFFECTIF	POIDS	MOYENNE	ECART-TYPE	MINIMUM	MAXIMUM
41 .	Fami - Famille, enfants : i	265	265.00	6.63	1.11	1.00	7.00
42 .	Trav - Travail, profession	265	265.00	5.98	1.55	1.00	7.00
43 .	Lois - Temps libre, détente	265	265.00	5.31	1.45	0.00	7.00
44 .	Amis - Amis, connaissances	265	265.00	5.17	1.40	1.00	7.00
45 .	Part - Parents, frères, soe	265	265.00	5.62	1.43	1.00	7.00
46 .	Reli - Religion : importanc	265	265.00	3.31	1.99	0.00	7.00
47 .	Poli - Politique, vie polit	265	265.00	3.13	1.82	0.00	7.00

Figura 3.5. Estadístiques descriptives (enquesta de referència)

STATISTIQUES SOMMAIRES DES VARIABLES CONTINUES							
EFFECTIF TOTAL :		50	POIDS TOTAL :		50.00		
NUM .	IDEN - LIBELLE	EFFECTIF	POIDS	MOYENNE	ECART-TYPE	MINIMUM	MAXIMUM
41 .	Fami - Famille, enfants : i	50	50.00	6.78	0.76	3.00	7.00
42 .	Trav - Travail, profession	50	50.00	5.80	1.52	1.00	7.00
43 .	Lois - Temps libre, détente	50	50.00	5.22	1.49	1.00	7.00
44 .	Amis - Amis, connaissances	50	50.00	5.32	1.52	2.00	7.00
45 .	Part - Parents, frères, soe	50	50.00	5.70	1.47	1.00	7.00
46 .	Reli - Religion : importanc	50	50.00	2.88	2.15	1.00	7.00
47 .	Poli - Politique, vie polit	50	50.00	3.02	1.48	1.00	6.00
28 .	Nbef - Nombre d'enfants eus	50	50.00	1.84	1.53	0.00	8.00
51 .	Salr - Salaire mens. de l'e	44	44.00	4842.05	5086.96	0.00	24000.00

Figura 3.6. Estadístiques descriptives (enquesta puntual)

TABLEAU DU DIFFERENCES DE CORRELATIONS

	Fami	Trav	Lois	Amis	Part	Reli	Poli
Fami	0.00						
Trav	-0.17	0.00					
Lois	0.10	0.13	0.00				
Amis	0.38	-0.22	0.13	0.00			
Part	0.19	-0.14	-0.09	0.17	0.00		
Reli	0.16	-0.21	-0.04	0.11	0.06	0.00	
Poli	0.18	0.17	0.14	0.18	0.13	-0.11	0.00

TABLEAU VTEST POUR LES CORRELATIONS

	Fami	Trav	Lois	Amis	Part	Reli	Poli
Fami	0.00						
Trav	-1.10	0.00					
Lois	0.67	0.86	0.00				
Amis	2.48	-1.44	1.01	0.00			
Part	1.35	-0.89	-0.58	1.16	0.00		
Reli	1.04	-1.36	-0.25	0.71	0.37	0.00	
Poli	1.15	1.07	0.91	1.17	0.81	-0.74	0.00

VALEURS-TEST DES MATRICES DES CORRELATIONS							
VARIABLE 1	VARIABLE 2	corr1	corr2	dif_corr	v_test	prob	
42. Trav	41. Fami	0.01	0.19	-0.17	-1.10	0.86	
43. Lois	41. Fami	0.20	0.10	0.10	0.67	0.25	
43. Lois	42. Trav	0.33	0.20	0.13	0.86	0.20	
...	...						
47. Poli	46. Reli	0.20	0.31	-0.11	-0.74	0.77	

Figura 3.7. Test de correlacions

COORDONNEES, CONTRIBUTIONS ET COSINUS CARRES DES INDIVIDUS SUR LES AXES 1 A 5

INDIVIDUS			COORDONNEES					CONTRIBUTIONS					COSINUS CARRES				
IDENTIFICATEUR	P.REL	DISTO	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1623	2.00	5.77	-1.30	-0.95	-0.08	-0.97	0.09	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00
1629	2.00	24.11	-1.45	1.70	-0.20	3.71	-1.41	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00
1635	2.00	-1.82	-0.04	-0.57	0.10	0.28	0.83	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00
1643	2.00	6.53	-1.38	-1.74	0.83	0.23	-0.36	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00
...																	
1898	2.00	6.13	1.87	-0.92	-0.87	0.04	0.59	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00
1904	2.00	2.99	0.46	-0.15	-0.13	0.20	1.03	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00
1910	2.00	6.72	0.75	1.67	1.48	-1.05	-0.02	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00
1916	2.00	4.61	1.51	0.79	1.09	-0.18	-0.19	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00
1922	2.00	4.93	0.12	0.18	-0.78	0.51	1.85	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00

Figura 3.8. Projectió en suplementari dels individus

CORRELATIONS ENTRE AXES DU NGUS ET NGUS1

Axes du NGUS1							
0.92	-0.29	0.03	-0.02	-0.04	-0.25	0.00	
0.14	0.52	0.74	-0.32	0.24	-0.03	-0.04	
0.08	0.12	0.36	0.87	-0.25	0.16	-0.09	
0.25	0.70	-0.56	0.16	0.26	-0.02	-0.21	
0.12	-0.29	0.03	-0.04	0.42	0.69	-0.49	
0.16	0.23	-0.08	-0.33	-0.76	0.46	-0.08	
0.16	0.06	-0.06	0.07	0.22	0.47	0.83	

COORDONNEES ET VALEURS-TEST DU CENTRE GRAVITE SUR LES AXES 1 A 5

	VALEURS-TEST					COORDONNEES					DISTO
	1	2	3	4	5	1	2	3	4	5	
Centre	-0.1	-1.0	-0.5	1.4	0.5	-0.01	-0.15	-0.07	0.20	0.08	0.10

Figura 3.9. Correlacions entre eixos i coordenades dels centres de gravetat

COORDONNEES DES VARIABLES SUR LES AXES 1 A 5
VARIABLES ACTIVES

VARIABLES	COORDONNEES					CORRELATIONS VARIABLE-FACTEUR					ANCIENS AXES UNITAIRES				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
IDEN - LIBELLE COURT															
Fami - Famille, enfants : i	0.52	-0.13	0.35	0.32	0.06	0.52	-0.13	0.35	0.32	0.06	0.00	0.00	0.00	0.00	0.00
Trav - Travail, profession	0.21	-0.49	0.23	-0.77	0.06	0.21	-0.49	0.23	-0.77	0.06	0.00	0.00	0.00	0.00	0.00
Lois - Temps libre, détente	0.69	-0.31	-0.49	-0.31	0.38	0.69	-0.31	-0.49	-0.31	0.38	0.00	0.00	0.00	0.00	0.00
Amis - Amis, connaissances	0.76	0.29	-0.55	0.28	-0.03	0.76	0.29	-0.55	0.28	-0.03	0.00	0.00	0.00	0.00	0.00
Part - Parents, frères, soe	0.61	0.05	0.11	0.68	-0.20	0.61	0.05	0.11	0.68	-0.20	0.00	0.00	0.00	0.00	0.00
Reli - Religion : importanc	0.25	0.80	0.39	0.15	-0.62	0.25	0.80	0.39	0.15	-0.62	0.00	0.00	0.00	0.00	0.00
Poli - Politique, vie polit	0.52	0.43	0.21	-0.15	0.45	0.52	0.43	0.21	-0.15	0.45	0.00	0.00	0.00	0.00	0.00

VARIABLES ILLUSTRATIVES

VARIABLES	COORDONNEES					CORRELATIONS VARIABLE-FACTEUR					ANCIENS AXES UNITAIRES				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
IDEN - LIBELLE COURT															
Nbef - Nombre d'enfants eus	0.10	0.16	0.33	0.14	-0.19	0.10	0.16	0.33	0.14	-0.19					
Salr - Salaire mens. de l'e	-0.12	-0.04	0.06	-0.10	-0.02	-0.12	-0.04	0.06	-0.10	-0.02					

Figura 3.10. Projectió doblement suplementària de les variables contínues comunes i no comunes

COORDONNEES ET VALEURS-TEST DES MODALITES
AXES 1 A 5

MODALITES			VALEURS-TEST					COORDONNEES					
IDEN - LIBELLE	EFF.	P.ABS	1	2	3	4	5	1	2	3	4	5	DISTO.
17. Disposez-vous d'une résidence secondaire													
Sec1 - oui	6	6.00	0.2	-0.8	-0.8	-0.6	0.7	0.13	-0.36	-0.32	-0.23	0.21	0.51
Sec2 - non	44	44.00	-0.4	-1.8	-0.6	5.1	1.4	-0.03	-0.12	-0.03	0.25	0.06	0.11
23. Avez-vous souffert récemment de nervosité													
Nrf1 - oui	26	26.00	-1.0	-1.8	0.1	1.7	1.0	-0.20	-0.30	0.01	0.23	0.11	0.22
Nrf2 - non	24	24.00	0.9	0.1	-1.0	1.1	0.3	0.20	0.02	-0.15	0.16	0.04	0.21

Figura 3.11. Projectió doblement suplementària de les variables categòriques no comunes

3.6.2 Interpretació dels resultats

Les conclusions més importants a què s'arriba a partir dels resultats numèrics són les següents:

- El test de correlacions (veure figura 3.7) informa que les característiques de les variables que conformen l'espai comú són semblants en ambdues enquestes, ja que només una parella de correlacions surt significativament diferent i que es considerarà casual (correlació entre importància a la família i importància a l'amistat, amb un valor-test de 2,48). La resta d'elements de les dues matrius són estadísticament equivalents. Observem també la forta correlació que existeix entre els dos primers vectors propis generats arran l'ACP de l'enquesta de referència i els generats per a l'enquesta puntual (0,92 i 0,52) (veure figura 3.9) mentre que per al tercer i quart les correlacions són més petites (0,36 i 0,16).

- Les variables comunes en l'enquesta puntual se situen en el primer eix factorial com a factor de creixement (veure figura 3.10), de la mateixa manera que s'havia vist que ho feien les variables de l'enquesta de referència. De la mateixa manera s'observa que les coordenades es diferencien de forma residual entre una enquesta i l'altra. Estaríem parlant llavors, d'un comportament molt semblant entre variables com ja s'ha comprovat via el test de correlacions.

- Pel que fa a les variables il·lustratives (aquelles que teòricament hem recollit en l'enquesta puntual i no ho havíem fet en l'enquesta de referència) s'observa (figura 3.10) que el nombre de fills té una correlació considerable (0,33) amb el tercer eix factorial (els enquestats amb més fills tendeixen a valorar menys l'oci i les amistat en contraposició a la resta de temes), mentre que el salari de l'enquestat té unes correlacions molt minses amb qualsevol dels 4 eixos.

- L'única modalitat de les variables categòriques significativa en els eixos factorials (veure figura 3.11) és en la què pertanyen els enquestats que no tenen segona residència, amb tendència a valorar per sobre de la resta la família i els amics.

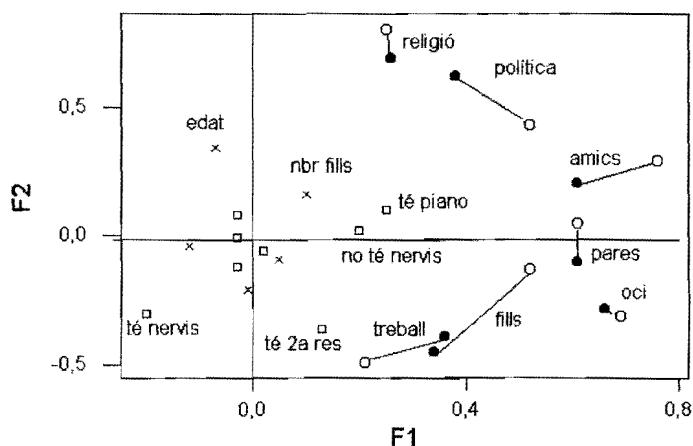


Figura 3.12. Representació de les variables actives de l'enquesta de referència (punts negres) comunes en l'enquesta puntual (punts blancs) i totes les variables il·lustratives en el pla factorial $\langle F_1, F_2 \rangle$.

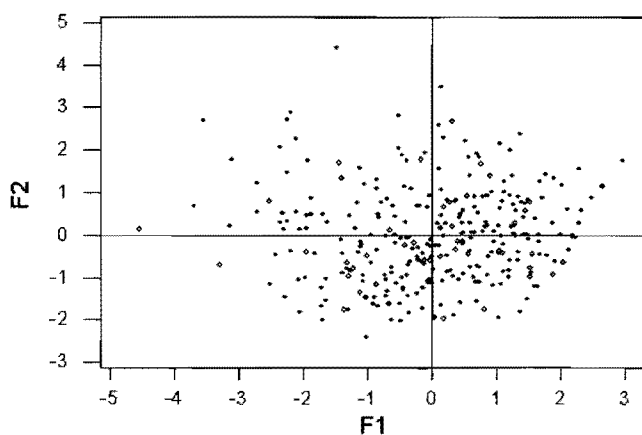


Figura 3.13. Representació dels individus enquestats en primera etapa (enquesta de referència – punts negres) i els enquestats en segona etapa (enquesta puntual – punts blancs) sobre el pla factorial $\langle F1, F2 \rangle$. S'observa la dispersió uniforme tant en un grup com en l'altre, quedant totalment barrejats els segons amb els primers, garantint distribucions de variables comunes molt semblants.

3.7 Una altra aplicació del procediment GREF1

Es procedirà a aplicar la inserció sobre unes altres dades on sí que seria necessari una translació per deixar constància d'una situació que es dona amb més freqüència en la realitat. Les dades estan generades per simulació. L'enquesta de referència està formada per 80 individus i té tres variables contínues, comunes a l'enquesta puntual, que té 50 individus. Tant una com l'altra enquesta tenen una variable categòrica il·lustrativa.

L'anàlisi factorial descriptiva de l'enquesta de referència mostra un valor propi molt significatiu (que conserva el 67% de la variabilitat de les dades) molt correlacionat amb les dues primeres variables contínues ($j1$ i $j2$) i que separa els individus amb un valor de variable molt alt en $j1$ i molt baix en $j2$ de la resta. El segon eix és molt menys significatiu i està molt correlacionat amb la variable $j3$. La variable categòrica és fortament significativa en el primer eix i la categoria M1 es relaciona amb els individus que tenen alt valor en $j2$ i baix en $j1$ i la categoria M2 amb la situació inversa.

En aplicar el procediment MIVAR sobre els resultats de l'ACP de l'enquesta de referència mantenint un sol eix com a significatiu, s'observa que eliminant la variable $j3$ de l'AFD es conserva més del 95% de la representativitat del subespai de representació, mentre que eliminant una de les altres dues variables, la pèrdua és més considerable. No obstant conservarem totes variables per a realitzar la inserció, ja que disposem d'elles en l'enquesta puntual.

De l'anàlisi dels resultats de la inserció (veure annex 3) destaquen les següents consideracions:

- L'estadística descriptiva de les variables mostra una mitjana molt diferent entre la variable j_1 en l'enquesta de referència i en l'enquesta puntual ($\overline{j1_0} = 0$ i $\overline{j1_1} = 2,05$) mentre que per a la resta de les variables les mitjanes coincideixen.

- Les correlacions entre les variables j_1 i j_2 en les dues enquestes són estadísticament diferents tot i que podrien semblar molt semblants en termes absoluts ($\text{corr}_0(j_1, j_2) = -0,91$ i $\text{corr}_1(j_1, j_2) = -0,96$), posant en perill la coherència global.

- Els vectors propis dels dos ACP tenen fortes correlacions (de l'ordre de 0,97), garantint la coherència en el subespai de representació.

- Les variables contínues de l'enquesta puntual s'insereixen amb coordenades molt semblants a les de l'enquesta de referència sobre el primer eix factorial de l'ACP, interpretant-se de forma equivalent.

- L'individu "centre de gravetat" creat a partir dels individus de l'enquesta puntual està significativament lluny de l'origen en tots els eixos que s'analitzen, concretament per al primer eix, té una coordenada positiva i amb un valor-test de 2,7. Això indicaria que el núvol puntual tendeix en conjunt a comportar-se diferent del de referència, tenint valors mes alts per a la variable j_1 i menys baixos alhora per a la variable j_2 . Això afecta a l'interpretació de les coordenades de les modalitats de l'enquesta puntual, situant-se la primera modalitat (MA) amb coordenada $-0,40$ i la segona modalitat (MB) en $1,04$, amb centre de gravetat en $0,32$, pròxim al del núvol puntual ($0,38$) i per tant, el valor-test per a MB resulta molt significatiu ($5,7$) però per a MA no ($-1,9$). La translació de les dades resoldria aquest problema ja que restaria a les coordenades dels individus i de les modalitats el centre de gravetat, i s'obtindria:

$\text{coord ("centre de gravetat")} \leftarrow 0,38 - 0,38 = 0 \rightarrow \text{valor-test} = 0 \text{ (no significatiu)}$

$\text{coord (MA)} \leftarrow -0,40 - 0,38 = -0,78 \rightarrow \text{valor-test} < -1,9 \text{ (significatiu)}$

$\text{coord (MB)} \leftarrow 1,04 - 0,38 = 0,66 \rightarrow \text{valor test} < 5,7 \text{ (significatiu)}$

- Gràficament, les situacions anteriorment descrites es representen de la següent manera:

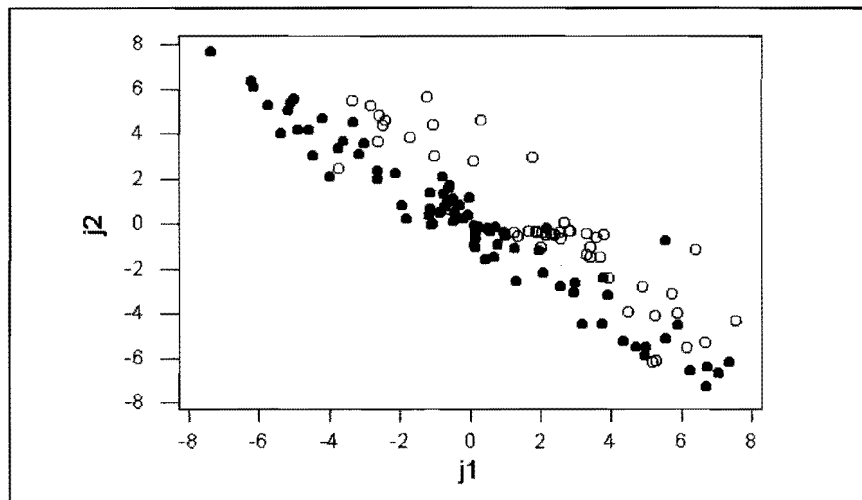


Figura 3.14. Representació dels individus de referència (punts negres) i els puntuals (punts blancs) en un sub-espai comú $\langle j1, j2 \rangle$ on es veu, a part de la forta correlació entre ambdues variables que el núvol puntual se situa per sobre del núvol de referència.

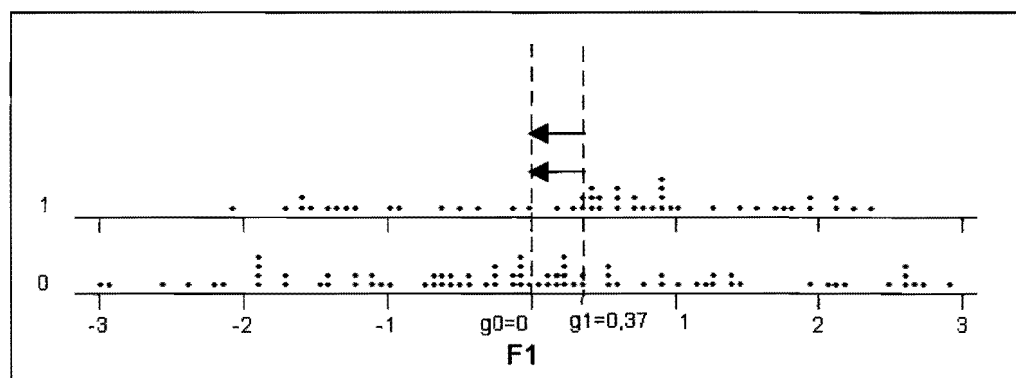


Figura 3.14. Representació dels individus de referència (0) en el primer eix factorial determinat per l'ACP i els individus puntuals inserits sobre el mateix eix, remarcant-ne els respectius centres de gravetat (g_0 i g_1).

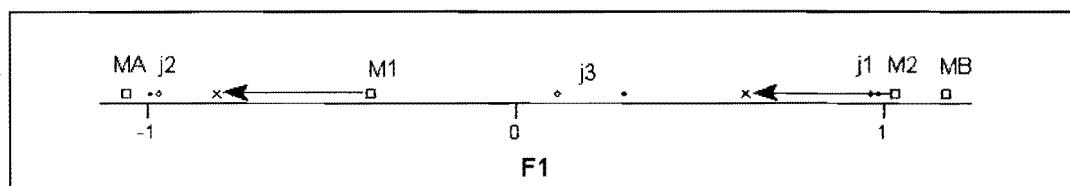


Figura 3.15. Representació de les variables de l'enquesta de referència actives en l'ACP (j_1 , j_2 , j_3 punt negre) comuns a les de l'enquesta puntual i inserides (j_1 , j_2 , j_3 punt blanc), les categories de la variable categòrica de l'enquesta de referència (MA i MB) i de l'enquesta puntual (M1, M2) amb la seva coordenada corresponent aplicant-hi la translació (x) sobre el primer eix factorial.

4. Un altre exemple complet

4.1 Presentació de les dades:

Per il·lustrar les capacitats dels dos procediments presentats en un exemple més real que els anteriors, basats en simulacions, es procedirà a analitzar dos conjunts de dades independents i a realitzar-ne la inserció. L'exemple es basa en relacionar dos estudis per a mesurar l'eficàcia d'una campanya publicitària. L'estudi anomenat de "referència" prové de SOFRES, empresa dedicada a controlar i mesurar l'audiència de televisió i realitza seguiments de publicitat, i és un pànel amb dades per a més de 2000 famílies sobre variables socio-econòmiques i d'audiència de les diferents emissores durant les franges horàries de cada dia a més a més del nombre d'impactes visionats de dues campanyes comercials. L'estudi anomenat "puntual" prové de la mateixa empresa i tracta sobre el consum dels espanyols que conté més de 4500 entrevistes amb dades socio-econòmiques, d'audiència i el consum d'un producte abans i després de les campanyes publicitàries de televisió.

Les variables comunes en ambdós estudis són les relacionades a dades socio-econòmiques (totes categòriques) i les que tracten l'audiència de les cadenes de televisió, constituïdes per tot un nombre de variables contínues que prenen valors des de 1 fins a 5 (de veure molt a veure gens respectivament) per a una cadena durant 6 franges horàries entre setmana, dissabte i diumenge.

4.2 Selecció d'individus i de variables

Els individus que se seleccionaran, per a simplificar l'exemple, seran aquells que pertanyen a la regió catalano-balear i que no tenen descodificador de canal plus, d'aquesta manera també es pot limitar a estudiar les variables d'audiència d'emissores que emeten en obert i en aquest territori (TVE1, TVE2, A3, T5, TV3 i C33). També es limitarà a estudiar el minutatge d'audiència entre setmana per a totes les franges horàries.

Les variables il·lustratives importants a relacionar entre estudis es prendran com a il·lustratives. En l'enquesta de referència són 4 variables contínues referents al nombre d'impactes publicitaris visionats per cada família per a 4 productes (caldos en general, caldos starlux, suavitzants en general i suavitzant flor). En l'enquesta puntual en seran 8, corresponents a la despesa en els 4 productes anteriors abans de la campanya i durant la campanya publicitària. A més a més s'agafaran dues variables categòriques com a il·lustratives també de cada enquesta : tamany familiar i tamany del municipi en l'enquesta de referència i

4.3. ACP de l'enquesta de referència

De la realització de l'ACP (no normalitzada) de l'enquesta de referència, destaquen les següents consideracions:

- Totes les variables tenen mitjanes molt altes, corresponents a que la tendència general a veure la televisió entre setmana és força baixa excepte la franja horària corresponent a la nit en que les mitjanes baixen per a totes les emissores. La variabilitat és força baixa (de l'ordre de la unitat) i les covariàncies són altes per a variables corresponents a franges horàries consecutives per a totes les cadenes, i també podrien ser significatives les covariàncies negatives que existeixen entre les emissores estatals i TV3, interpretables com que hi ha força discrepància entre espectadors de les primeres i la segona emissora.

- Hi ha un primer eix factorial que destaca, amb una inèrcia del 21% que contraposa els espectadors que veuen les emissores d'àmbit nacional amb els espectadors de TV3 doncs les variables relacionades amb aquesta cadena estan en negatiu mentre que la resta estan en positiu.

- El segon eix, amb un 12% de la inèrcia total, s'interpreta com un factor de creixement, amb valor negatiu per a totes les variables excepte dues (espectadors de TVE1 per la tarda).

- El tercer eix factorial, menys representatiu (9% de la inèrcia) sembla diferenciar els individus que veuen la televisió pel matí dels qui ho fan la resta del dia.

- Les variables il·lustratives (spots visionats) tenen coordenades negatives al primer eix (indicant que els espectadors de les emissores nacionals veuen més anuncis) i positives en el segon eix, com és lògic, doncs els qui miren més la tele, veuen més anuncis. Per que fa a les variables categòriques escollides, només sembla significatiu el pertànyer en pobles més petits on es veu més TV3 que la resta de cadenes.

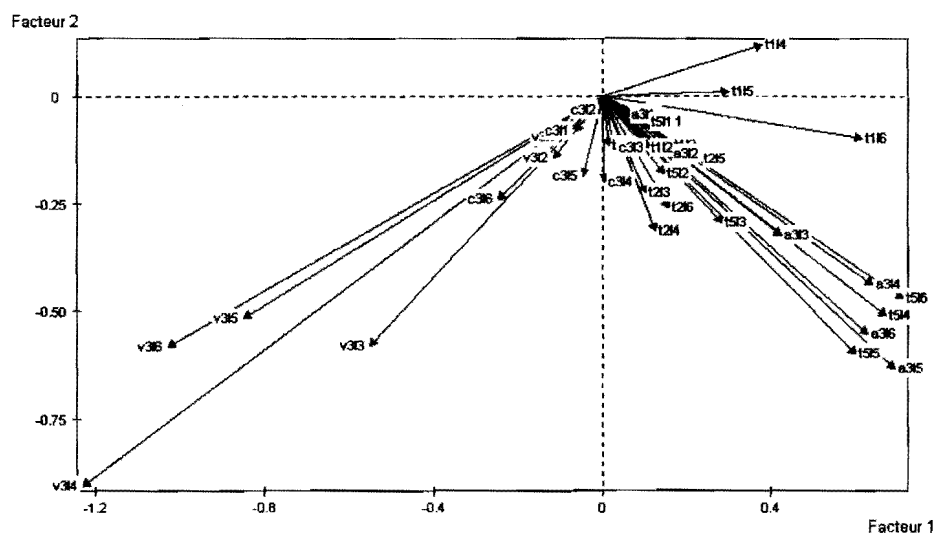


Figura 4.1. Representació de les variables actives sobre el pla factorial obtingut amb l'ACP de l'enquesta de referència

4.4 Selecció del subespai de representació

Per a seleccionar el marc de representació o subconjunt de variables que conformen el millor subespai de representació s'utilitzarà el procediment MIVAR, calculant el criteri de selecció per a 2 eixos factorials i conservant un mínim de 12 variables de les 36 que formen part de l'enquesta de referència. Els resultats obtinguts són els següents:

- El criteri inicial (suma dels valors propis) és 11,589.

- Conservant 12 variables (una tercera part del total) el criteri només disminueix en una unitat i escaig (un 11%), donant la idea que totes les variables estan molt correlacionades i que per a relacionar les variables d'una enquesta variable no serà necessari ni molt menys conservar tota l'estructura de dades.

- Les variables que primer es rebutgen són les relacionades amb l'audiència durant el matí per a totes les cadenes (durant les 3 primeres franges horàries excepte per A3, T5 i TV3 on la 3ª franja és important). Les variables que corresponen al C33 surten totes a partir de la 24 passada del procediment, i les de TVE2 a partir de la 25, "sobrevivint" en el subconjunt mínim: la franja de nit de TVE1, A3, T5 i TV3, les franges de tarda de A3, T5 i TV3 i la franja de migdia de A3 i TV3.

- Els resultats anteriors semblen lògics doncs durant la nit és quan es concentren més telespectadors en totes les emissores i les correlacions que existeixin entre variables quedaran més patents durant aquestes franges que en la resta del dia. A més a més és durant aquest horari quan més impactes publicitaris rebran les famílies i per tant, relacionar les variables il·lustratives també amb aquestes variables sembla el més adequat.

- Com a implicació dels resultats del procediment MIVAR es procedirà a definir el sub-espai de representació. Les variables considerades comunes en la inserció seran les 12 determinades pel subconjunt mínim òptim.

- Efectivament, realitzant una ACP mantenint les 12 variables com a actives, els resultats obtinguts són idèntics als que s'obtenen amb 36 variables: primer eix que separa espectadors de TV3 de la resta de cadenes, segon factor, de creixement. Les variables il·lustratives mantenen una relació amb les actives molt similar al primer ACP. Amb aquest ACP s'ha creat el que serà el subespai de representació on s'inseriran les dades puntuals. En el següent apartat es realitza l'operació.

4.5 Inserció de dades

S'ha realitzat una inserció de les dades puntuals al subespai de representació determinat per les 12 variables contínues comunes (veure Annex4) i els resultats més destacables són els següents:

- L'estadística descriptiva de les variables comunes dona mitjanes i desviacions típus molt semblants en ambdues enquestes i en principi els individus de l'enquesta puntual en mitjana veuen les cadenes durant les franges horàries corresponents en la mateixa mesura que els de l'enquesta de referència.

- Les correlacions entre els dos vectors propis originals d'ambdós anàlisis són molt altes (0,91 i 0,92) i més baixes entre la resta de vectors (no significatius per a l'anàlisi).

- El test de correlacions ofereix unes diferències força significatives entre correlacions de variables en ambdues enquestes. Per exemple, les variables a315 i a316 presenten una correlació en l'enquesta puntual 0,5 unitats més gran que en l'enquesta de referència. Això estaria indicant algun problema de coherència global doncs les característiques d'interrelació de variables no són les mateixes en els dos estudis.

- Les variables inserides de l'enquesta puntual comunes a l'enquesta de referència tenen coordenades sobre els dos primers factors del mateix signe que en l'ACP de referència (el primer eix separa els espectadors de TV3 dels de la resta de cadenes i el segon separa els espectadors que en general miren més la tele que els que no la miren gaire).

- Per què fa a les variables contínues il·lustratives, relatives al consum de 4 productes abans i durant la campanya televisiva, tenen coordenades negatives per al primer eix, igual que les variables de l'enquesta de referència referents al nombre d'espots visionats, per tant sembla existir correlació entre veure més anuncis (en les cadenes nacionals) i comprar els productes, tant abans com durant la campanya. En el segon factor, de creixement, les variables de visionat d'anuncis (en els 4 productes) es relaciona amb la compra de suavitzant, mentre que per a la compra del suavitzant flor i de pastilles de caldo, no ho està. En resum, sembla que la campanya publicitària té efecte sobretot en la compra de suavitzant, variable més correlacionada, sobretot durant la campanya, a la de nombre d'espots visionats. Pel que fa al suavitzant flor en concret, no existeix la relació i de la mateixa manera per les pastilles de caldo, independents a les campanyes publicitàries corresponents.

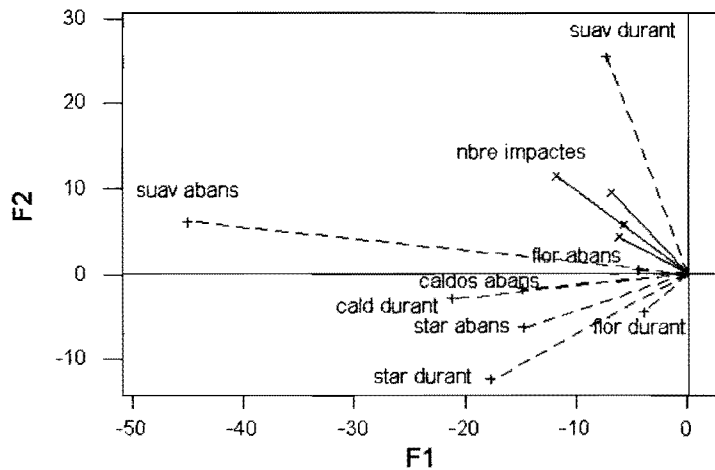


Figura 4.2. Pla factorial $\langle F1, F2 \rangle$ amb la representació de les variables contínues il·lustratives de referència (línia contínua) i puntuals (línia discontinua).

· Analitzant les coordenades dels individus, cal remarcar que el centre de gravetat en ambdues enquestes són significativament diferents, doncs el núvol puntual se situa 7 i 21 centèssimes per la dreta i per l'esquerra respectivament dels dos primers eixos, amb uns valors-test de 2,2 i -6 unitats. Això ve a dir que el segon grup en estudi està en mitjana molt per sota en el segon eix (la distància en el primer eix es considera residual), indicant que veuen menys la televisió que els enquestats en primera etapa. Aquest resultat distorsiona els resultats per a les categories il·lustratives.

· Les variables categòriques “possessió d’aspirador” i “de cotxe” es col·loquen en el primer eix de forma significativa, i els que tenen aspirador i més d’un cotxe veuen més TV3 que els que no tenen menys equipament que tendeixen a veure les cadenes nacionals. La possessió d’aspirador i cotxe pot ser una mesura primària del nivell socio-econòmic de la família i segons aquesta interpretació, les classes més altes veuen el canal autonòmic amb més freqüència que les baixes. Les categories en el segon eix, havent modificat les coordenades pel centre de gravetat dels individus que les conformen i els valors-test corresponents no resulten significatives.

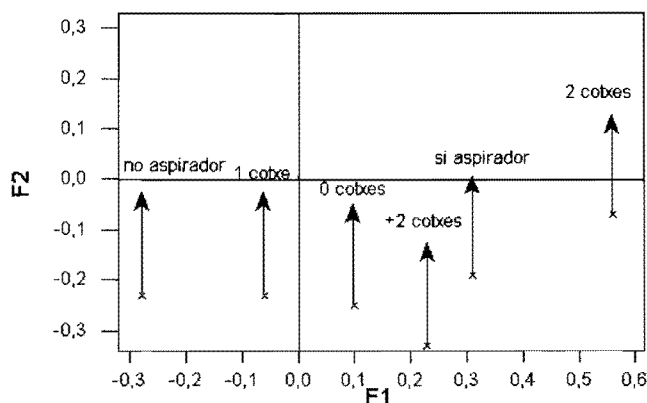


Figura 4.3. Categories il·lustratives puntuals traslladades en el pla factorial $\langle F1, F2 \rangle$

4.6 Conclusions

- Un cop realitzada la inserció podem concloure que les variables que millor conserven la variabilitat dels individus és l'audiència per la nit en totes les cadenes excepte en TV2 i C33 precisament perquè són les variables amb més desviació tipus i en una anàlisi no normalitzada, pesa molt la variabilitat de les variables a l'hora de formar els eixos.

- Inserint les variables puntuals veiem que el comportament del pànel de consum és molt similar al del pànel d'audiència i que les condicions per inserir són bones, encara que el test de correlacions ha informat de les considerables diferències que existeixen entre les estructures d'interrelacions de les variables en les dues enquestes.

- Les variables il·lustratives no han mostrat una relació gaire òptima des del punt de vista de l'efectivitat del màrketing ja que sembla que l'únic consum susceptible a campanyes de publicitat sigui el de suavitzants en general, mentre que per a la marca flor i per a les pastilles de caldo, sembla independent al nombre d'espots que les famílies vegin sobre ells.

5. Conclusions

El projecte ha versat sobre l'actualització de dues rutines dissenyades en el seu moment per aplicar la teoria de la inserció de dades dins del paquet estadístic SPAD. Això a permès en primera instància descobrir l'eina en qüestió i poder aplicar-la i de la mateixa manera, aplicar els meus coneixements de programació en un problema real.

Per a introduir la inserció de dades d'enquesta s'han definit les característiques de les AFD, i l'ACP en concret, perquè la inserció funciona a partir d'una d'aquestes anàlisis, així com les sortides que genera SPAD per a aquests mètodes. S'ha explicat també en què consisteix la inserció, els càlculs que realitza i algunes eines suplementàries com pot ser el propi MIVAR, que selecciona el subconjunt mínim de variables per formar el subespai de representació, o la translació i la rotació.

Tot i que el disseny dels programes ja havia estat en el seu moment proposat, decidit i programat, la reconversió a permès entrar en el camp de la programació pràctica, aquella que acaba basant-se en una baralla amb el compilador per tal de anar restant errors i amb controls en el codi per saber en quina instrucció es queda encallada l'execució i la programació comercial, basada en un sistema que ja existeix, que utilitza un llenguatge determinat i unes regles estrictes a seguir.

Cal destacar també la feina realitzada per a millorar la sortida en pantalla del procediment MIVAR, resumint-la en una taula i aportant-hi resultats gràfics, de la mateixa manera que la intenció de dissenyar per al procediment GREF1 una subrutina que genera a partir d'una base, dos fitxers diccionaris de dades, dos fitxers de coordenades i un fitxer amb les coordenades que genera l'ACP sobre un dels dos conjunts de dades, imprescindible per a utilitzar el procediment en un entorn Windows d'SPAD, avortada per manca de temps. Per a la realització de les anteriors actualitzacions ha estat clau la experiència en aquest camp de la Roser Rius i l'exemple de les rutines definides per a les últimes versions d'SPAD, incloses en el seu programari.

Les diferències bàsiques que existeixen entre els programes dissenyats per a les versions en desenvolupament d'SPAD anteriors a 4.0 i les que funcionen en 32 bits són: l'administració de memòria, la definició explícita de variables, l'ús d'un sol fitxer '.for' on s'inclouen totes les subrutines que componen el programa i algunes qüestions de format.

L'aplicació dels dos procediments ha confirmat la comprensió de les metodologies que segueixen i la oportunitat d'analitzar-ne els resultats que se'n desprenen a més a més de la utilitat realment important que tenen en molts camps d'actualitat indiscutible, com pot ser en estudis d'audiència televisiva (camp per a les quals van ser desenvolupades).

Cal tenir en compte que només s'ha actualitzat el procediment GREF1 per a dades originàries d'una anàlisi en components principals, sense translació ni rotació i per tant, un treball futur que queda pendent és la implementació d'aquests dos sub-programes així com la inserció de dades provinents d'anàlisis en correspondències i la implementació de la inserció conjunta.

Annex1. Programa MIVAR complet

```

C+++++ FDDUM1
++
C+ -----
C+ MIVAR
+
C+ -----
C+ DUMM1, CPVAR, SOUSDI, SOUSDO, CORIN2, CORAXE, CALCRI.
C+++++
SUBROUTINE DUMM1 (NXMQ,NXKQ,NXRQ,NXDQ,LISTF,LRESM,LERRE,LEPAR,
& KFICH,NFICH,LEXCE)
C*****
**C* MIVAR
C* PARAMETRES LUS SUR LEC
C* -----
C* NAXR..... NOMBRE D'AXES _ RECONSTITUER
C* NMVAR..... NOMBRE MINIMUM DE VARS POUR LE SOUS-ENSEMBLE
C* LEDIN..... IMPRESSION DES COORDONNEES DES INDIVIDUS
C* 0 - PAS D'IMPRESSION
C* 1 - IMPRESSION POUR LES INDIVIDUS ACTIFS
C* 2 - IMPRESSION POUR TOUS LES INDIVIDUS
C*
C* APPELS DUMM1 -- DEBET *
C* -----
C* -- ONDIC ... *
C* -- ONGUS ... *
C* -- PARAM ... *
C* -- CPVAR -- OFICH *
C* -- SOUSDI -- ONDIC ... *
C* -- SOUSDO -- ONDON ... *
C* -- ONDON ... * -- ONDIC ... *
C* -- STABA ... *
C* -- COREL ... *
C* -- VPROP ... *
C* -- CORIN2 *
C* -- CORAXE *
C* -- EDCOR ... *
C* -- CALCRI *
C* -- SHELK *
C* -- VANEG *
C* -- FINET *
C*
C* FICHIERS EN ENTREE .... NDIC (DICTIONNAIRE UTILE)
C* ----- NGUS (COORDONNEES FACTORIELLES)
C*
C* DE TRAVAIL ... NSAV (DONNEES CENTREES ET REDUITES)
C* ----- NBAND (SAUVEGARDES DIVERSES)
C*
C*****
C***** DECLARATIONS EXPLICITES ET DIMENSIONNEMENT
INTEGER NXMQ, NXKQ, NXRQ, NXDQ, LISTF, LRESM, LERRE, LEPAR,
& NFICH, LEXCE
INTEGER LEC, ILLEC, IMP, NXCPA, NXLPA, NLIMP
INTEGER NDICZ, NDONZ, NDICA, NDONA, NDICB, NDONB, NDIC, NDON,
& NGRI, NGRO, NSAV, NBAND, NBFOR, NGUS,
INTEGER MDICZ, MDONZ, MGUSZ, MGUSG, NMACZ, NMACA, NTRA, NTRA2,
INTEGER FICLOG, NBASE
INTEGER NCOEF, NCOEFB
INTEGER NVTOT, NXMOD, NXMAC, NMTOT, NMACT, NITOT, NIACT,
& LANAL
INTEGER NQTOT, NQACT, NCONT, NCACT, NFTOT
INTEGER NAXE, LORIG, NIDI, NIDV
INTEGER MPARM(2)
INTEGER NAXR, NMVAR, LEDIN
INTEGER NUMREG, L CORR, NFI, NPOSIB, ND, NS, NE, NL,
& IERR
INTEGER invsta, iisous, iorjva, inbmod, iorin, iisum,
& imatv, ikidm, ikvar, ikmod, ikidi, ikidv,
& ikidax, ivar, ictn, ievr, iptot, icmoy,
& icvar, ivmin, ivmax, icorr, ivmoy, ivalpr, icind,
& itcorr, iaxes
REAL PITOT, PIAC, TEST
REAL VBI(1)
CHARACTER*4 KETA1, KETA2
CHARACTER*4 KTITR(20)
CHARACTER*256 KFICH (NFICH)
C=====
INCLUDE 'DINTAB.INC'
COMMON /ENSOR/ LEC, ILLEC, IMP, NXCPA, NXLPA, NLIMP
COMMON /NUFIC/ NDICZ, NDONZ, NDICA, NDONA, NDICB, NDONB,
& NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR
COMMON /NOFIC/ MDICZ, MDONZ, MGUSZ, MGUSG, NMACZ, NMACA,
& NTRA, NTRA2, NBASE
COMMON /ILOG/ FICLOG
COMMON /NDISC/ NCOEF, NCOEFB

```

```

DATA KETA1 /'DUMM'/, KETA2 /'1'/
C===== ALLOCATION POUR LECTURE DES PARAMETRES =====
ALLOCATE (MQ(NXMQ),STAT=LERRE)
IF (LERRE .NE. 0) GOTO 901
ALLOCATE (KQ(NXKQ),STAT=LERRE)
IF (LERRE .NE. 0) GOTO 901
ALLOCATE (RQ(NXRQ),STAT=LERRE)
IF (LERRE .NE. 0) GOTO 901
ALLOCATE (DQ(NXDQ),STAT=LERRE)
IF (LERRE .NE. 0) GOTO 901
C=====
DQ(1) = 0.0
CALL DEBET (KETA1,KETA2,KTITR)
C===== LECTURE DES PARAMETRES SUR NDIC =====
CALL ONDIC(KQ,NVTOT,NXMOD,NXMAC,MQ,NMTOT,NMACT,NITOT,
& NIACT, PITOT, PIAC, LANAL, 1, LISTF, LERRE, NDIC)
IF (LERRE .NE. 0) GOTO 40
NQTOT = MQ(1) + MQ(2)
NQACT = MQ(1)
NCONT = MQ(3) + MQ(4)
NCACT = MQ(3)
NFTOT = MQ(5) + MQ(6)
C===== LECTURE DES PARAMETRES SUR NGUS =====
CALL ONGUS (KQ, NAXE, LORIG, NIDI, NIDV, MQ,
& 1, LISTF, LERRE, NGUS, '2')
IF (LERRE.NE.0) GOTO 40
C===== LECTURE DE PARAMETRES SUR LEC =====
MPARM(1) = NAXE
MPARM(2) = NCACT - 1
CALL PARAM (MQ,NXMQ, KQ,NXKQ, RQ,NXRQ, MPARM,2, VBI(1),
& 0, NXDQ, KETA1, LEPAR, LRESM, LERRE)
IF (LERRE .NE. 0) GOTO 40
NAXR = MIN0(MQ(1),NAXE)
NMVAR = MQ(2)
LEDIN = MQ(3)
LEXCE = 0
IF (NAXR.GT.NMVAR) WRITE(IMP,1000) NAXR
C===== RESERVATION POUR CPAR =====
NUMREG = 1
LCORR = 2 - LORIG
NFI = MAX0(NAXR+1,NMVAR+1)
NPOSIB = NCACT
ND = NCACT + 1
NS = NUMREG + 1
NE = ND * (ND + 1) / 2
NL = ND * (ND + 1) * (ND + 2) / 6
C===== RESERVATION DE ENTIERES =====
invsta = 1
iisous = invsta + 8
iorjva = iisous + (NCACT + 1) * NPOSIB
inbmod = iorjva + NVTOT
iorin = inbmod + NVTOT
iisum = iorin + NITOT
imatv = iisum + NCACT
nmq = imatv + NCACT * NCACT
C===== RESERVATION DE CARACTERS =====
ikidm = 1
ikvar = ikidm + NXMOD
ikmod = ikvar + 15
ikidi = ikmod + 5 * NXMOD
ikidv = ikidi + NIDI
ikidax = ikidv + NCACT
nkq = ikidax + NAXE
C===== RESERVATION DE REELS =====
ivar = 1
ictn = ivar + NQTOT
ievar = ivar + NVTOT
iptot = ievar + NCONT
icmoy = iptot + NCONT
icvar = icmoy + NCONT
ivmin = icvar + NCONT
ivmax = ivmin + NCONT
icorr = ivmax + NCONT
ivmoy = icorr + NCACT * NCACT
ivalpr = ivmoy + NCACT
icind = ivalpr + NCACT
iaxes = icind + NCACT
itcorr = iaxes + NIACT * NAXE
nrq = itcorr + (3 + NAXR + NAXR) * NCACT
C===== RESERVATION DE DOUBLE PRECISION =====
ndq = 1
C===== DEALLOCATION APRES LECTURE DES PARAMETRES =====
DEALLOCATE (MQ, STAT=IERR)
DEALLOCATE (KQ, STAT=IERR)
DEALLOCATE (RQ, STAT=IERR)
DEALLOCATE (DQ, STAT=IERR)
C===== ALLOCATION DYNAMIQUE DES TABLEAUX POUR EXECUTION =====
ALLOCATE (MQ(NMQ), STAT=LERRE)

```

```

IF (LERRE .NE. 0) GOTO 902
ALLOCATE (DQ(NDQ), STAT=LERRE)
IF (LERRE .NE. 0) GOTO 902

C ===== APPEL DE LA GESTION DE LA PROCEDURE =====

CALL CPVAR (NVTOT, NCACT, NCONT, NXMOD, NIDI, NAXE, NAXR, NFI,
&          NIACT, TEST,
&          NPOSIB, PIAC, NITOT, LCCORR, LEDIN, NQTOT,
&          MQ(invsta), MQ(iisous), MQ(ijorva),
MQ(inbmod),
&          MQ(ijorin), MQ(iisum), MQ(iimativ),
&          KQ(ikidm), KQ(ikvar), KQ(ikmod), KQ(ikidi), KQ(ikidv),
&          KTITR, KQ(ikidax),
&          RQ(ivar), RQ(ictn), RQ(ievar), RQ(iptot),
RQ(icmoy),
&          RQ(icvar), RQ(ivmin), RQ(ivmax), RQ(icorr),
RQ(ivmoy),
&          RQ(ivalpr), RQ(icind), RQ(iaxes),
RQ(itcorr), LEXCE)

C ===== FIN DE LA PROCEDURE =====
40 CALL FINET (KETAL, KETA2, KTITR)
IF (LEXCE .EQ. 1) THEN
WRITE (NCOEF, 5010)
WRITE (NCOEF, 5020)
ENDIF
GOTO 999

C ===== ERREUR D'ALLOCATION =====
901 CALL MALLO (KETAL, KETA2, NXMQ, NXKQ, NXRQ, NXDQ)
GOTO 40
902 CALL MALLO (KETAL, KETA2, NMQ, NKQ, NRQ, NDQ)
GOTO 40

C ===== DEALLOCATION DES TABLEAUX APRES EXECUTION (SI LERRE = 0) =====
999 IF (LERRE .EQ. 0) THEN
DEALLOCATE (MQ, STAT=IERR)
DEALLOCATE (KQ, STAT=IERR)
DEALLOCATE (RQ, STAT=IERR)
DEALLOCATE (DQ, STAT=IERR)
ENDIF
RETURN

C ===== FORMATS =====
1000 FORMAT (
& 'X ', '/', ' ', 'ATENCIO:EL SUJUNT DE VARIABLES MES PETIT ES DE
& ', I4, ' VARIABLES.', '/',
& 'X ')
5010 format ('ZZF')
5020 format('EOF')
9200 FORMAT (
FE & 'X ', '/', ' ', 'ERREUR (COPRI-920)', '/',
FE & ' ', 'ON TROUVE DES REPONSES MANQUANTES POUR CERTAINES ',
FE & ' ', 'VARIABLES NOMINALES.', '/',
FE & ' ', 'VOUS DEVEZ SPECIFIER "LZERO=REC" DANS LA ',
FE & ' ', 'PROCEDURE "SELEC" QUI PRECEDE.', '/',
FE & 'X ')
FE
END

SUBROUTINE CPVAR (NVTOT, NCACT, NCONT, NXMOD, NIDI, NAXE, NAXR,
&          NFI, NIACT, TEST, NPOSIB, PIAC, NITOT,
&          LCCORR, LEDIN, NQTOT,
&          nvsta, isous, jorva, nbmod, jorin, isum,
imativ,
&          kidm, kvar, kmod, kidi, kidv, ktitr, kidax,
&          var, ctn, evar, ptot, cmoy, cvar, vmin, vmax,
corr,
&          vmoy, valpr, cind, axes, tcorr, LEXCE)

C ===== DECLARATIONS EXPLICITES ET DIMENSIONNEMENT =====
INTEGER NVTOT, NCACT, NCONT, NXMOD, NIDI, NAXE, NAXR, NFI,
&          NIACT, NPOSIB, NITOT, LCCORR, LEDIN, IMP, NQTOT
INTEGER NDICZ, NDONZ, NDICA, NDONA, NDICB, NDONB, NDIC,
&          NDON, NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR, LEC,
&          ILLEC, NXCPA, NXLPA, NLIMP, IC, NV, MAXVAR, ICAS
INTEGER nvsta(8), isous(NCACT+1, NPOSIB), jorva(NVTOT),
&          nbmod(NVTOT), jorin(NITOT), isum(NCACT),
&          imativ(NCACT, NCACT)
INTEGER i, j, l, nb, NPASS, NBID
INTEGER NVARNO, NVTOT, NXMODP, NXMACP, NMTOTP, NMACTP, NITOTP,
&          NIACTP, LANAL, LERRE, NQEXAP, NQTOTP, NCONTP,
NCACTP,
&          NFTOTP, NISUP, NISUPP, nuls, NSUP
INTEGER LEXCE, FICLOG, NETOI, NBLAN
CHARACTER*4 kidm(NXMOD), kvar(15), kmod(5, NXMOD), kidi(NIDI),
&          kidv(NCACT), ktitr(20),
&          kidax(NAXE)
CHARACTER*4 KETOI
CHARACTER*4 KBLAN
CHARACTER*4 KIKI
CHARACTER*4 klib(:, :), KLIB0(5), kcon(15)
REAL TEST, PIAC,
REAL PNUL, ZERO, RMIN, RMAX, P, D, tcrit, TCORMAX, VANEG
REAL var(NVTOT), ctn(NCONT), evar(NCONT), ptot(NCONT),
&          cmoy(NCONT), cvar(NCONT), vmin(NCONT), vmax(NCONT),
&          corr(NCACT, NCACT), vmoy(NCACT), valpr(NCACT),
cind(NAXE),
&          axes(NIACT, NAXE),
&          tcorr(3+2*NAXR, NCACT)
REAL PITOTP, PIACTP, TRAC1, TRAC2, DOMAX, ECHEL

C =====
COMMON /SEUIL/ ZERO, RMIN, RMAX
COMMON /NUFIC/ NDICZ, NDONZ, NDICA, NDONA, NDICB, NDONB, NDIC,
&          NDON, NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR
COMMON /ENSOR/ LEC, ILLEC, IMP, NXCPA, NXLPA, NLIMP
COMMON /ILOG/ FICLOG
DATA KETOI /'*/', KBLAN /' '/

C ===== INITIALISATIONS =====
PNUL = VANEG (PIACT)
DO 10 i = 1, NCACT
isum(i) = 0
DO 10 j = 1, NCACT
corr(i, j) = 0
imatv(i, j) = 0
10 CONTINUE

C ===== LECTURE valpr SUR NGUS =====
CALL OFICH (2, NGUS)
READ(NGUS) (valpr(i), i = 1, NAXE)

C ===== ECRITURE INITIAL =====
tcrit = 0
DO 15 i = 1, NAXR
tcrit = tcrit + valpr(i)
15 CONTINUE
WRITE (IMP, 3000)
WRITE (IMP, 800) NAXR
WRITE (IMP, 3000)
WRITE (IMP, 801) NCACT, tcrit
WRITE (IMP, 3000)
WRITE (IMP, 802) NAXR
DO 16 i = 1, NAXR
16 WRITE (IMP, *) valpr(i)
WRITE (IMP, 3000)
WRITE (IMP, 803) NCACT

C ===== LECTURE LIBELLES SUR NDIC =====
NPASS = 3 + NQTOT
CALL OFICH (NPASS, NDIC)
DO 111 i = 1, NCACT
READ (NDIC) nbid, KIKI, (kcon(J), J=1, 15),
- (KLIB0(J), J=1, 5)
WRITE (IMP, *) ' ', i, ': ', KIKI, ' ', (KCON(j), j=1, 15)
111 CONTINUE
WRITE (IMP, 3000)

C ===== LECTURE DES AXES SUR NGUS =====
DO 20 IC = 1, NIACT
READ(NGUS) (axes(IC, i), i=1, NAXE), P, D
20 CONTINUE

C ===== INITIALISATION DE isous =====
DO 25 i = 1, NCACT + 1
DO 25 j = 1, NCACT
IF (j .EQ. i) THEN
isous(i, j) = 0
ELSE
isous(i, j) = 1
ENDIF
25 CONTINUE

C ===== ECRITURE AVANT BUCLE ELIMINATION =====
WRITE (IMP, 3000)
WRITE (IMP, 890) NFI-1

C ===== GRAND-GRAND BUCLE POUR CHAQUE ELIMINATION =====
49 DO 50 NV = NCACT, NFI, -1
TCORMAX = -RMAX
MAXVAR = 0
DO 90 i = 1, NCACT
tcorr(1, i) = 0.
tcorr(2, i) = i
tcorr(3, i) = 0.
DO 95 j = 4, 3+2*NAXR
tcorr(j, i) = 0.
95 CONTINUE
90 CONTINUE

C ===== GRAND BUCLE POUR CHAQUE SOUSSENSAMBLE =====
DO 100 ICAS = 1, NPOSIB
IF (isous(NCACT+1, ICAS) .GE. NCACT) GOTO 100

C ===== CHERCHER LE SOUSSENSAMBLE DANS NDIC ET NDON =====
NVARNO = isous(NCACT+1, ICAS)
CALL SOUSDI (NDIC, NSAV, isous(1, ICAS), NCACT, nvsta, 3,

```

```

&          NVARNO, 0, nbmod, jorva, NVTOT, jorin, NITOT,
kidm,
&          kvar, kmod, NXMOD, ktitr)
CALL SOUSDO (NDON, NBAND, isous(1,ICAS), NCACT, nvsta, 3,
&          NVARNO, var, NVTOT, kidi, NIDI, ktitr)

C ===== NOVEAUX FICHIERS: NDIC -> NSAV, NDON -> NBAND =====
C ===== CENTRAGE ET REDUCTION =====
CALL ONDIC (ktitr,NVTOT,NXMODP,NXMACP,nvsta,NMTOT,NMACTP,
&          NITOTP, NIACTP, PITOTP, PIACPT, LANAL, 1, 0,
LERRE,
&          NSAV)
CALL ONDON (ktitr,NVTOT,NIDI,NITOT,TEST,1,0,LERRE,NBAND)
NQEXAP = NVTOT - 1
NQTOTP = nvsta(1) + nvsta(2)
NCONTP = nvsta(3) + nvsta(4)
NCACTP = nvsta(3)
NFTOTP = nvsta(5) + nvsta(6)
ALLOCATE (klib(15,NCONTP), STAT = LERRE)
IF (LERRE.NE. 0) THEN
  WRITE(FICLOG,9000)
  RETURN
ENDIF
READ(NSAV) (nb, l=1, NVTOT+NQTOTP), (jorva(j), j=1, NCONTP)
CALL STABA (evar, ptot, cmoy, cvar, vmin, vmax, ctn, jorva,
&          NCONTP, var, NQEXAP, kidv, NCACTP, NQTOTP,
NIACTP,
&          PIACPT, TEST, PNUL, LERRE, NBAND, NSAV, 0,
NCACT,
&          klib, LEXCE)
DEALLOCATE (klib)
IF(LERRE.NE. 0) RETURN
NISUP = NITOT - NIACT
NISUPP = NITOTP - NIACTP
CALL COREL (ptot, cmoy, cvar, ctn, NCONTP, var, NQEXAP, corr,
&          vmoy, NCACTP, kidi, NIDI, NIACTP, NISUP, PIACPT,
PNUL,
&          TEST, LECORR, NBAND, NSAV, 0)
C ===== DIAGONALISATION =====
CALL VPROP (corr, valpr, ctn, NCACTP, NCACTP, TRAC1, TRAC2, LERRE)
C ===== CALCUL DES CORDONEES DES INDIVIDUS =====
REWIND NBAND
CALL CORIN2(NAXR, NCONTP, NCACTP, NQEXAP, NIDI, NITOTP, valpr,
&          kidi, var, cind, corr, ctn, DOMAX, NSAV, NBAND)
C == CORRELATION AXES: INITIALS(axes01),
axes_sousensamble(axes02) =
REWIND NBAND
CALL OFICH(3, NGUS)
CALL CORAXE (corr, ptot, cind, vmin, cvar, vmax,
&          NAXR, NIACTP, PNUL, TEST, NBAND, NGUS)

C === VERIFIER LE CAS MOINS BON POUR SAVOIR LA VAR A ELIMINER ===
510 CALL CALCRI (valpr, NAXR, corr, tcorr, 3+2*NAXR, NCACT,
ICAS)
IF(tcorr(1,ICAS) .GT. TCORMAX) THEN
  TCORMAX = tcorr(1,ICAS)
  MAXVAR = ICAS
ENDIF
GOTO 100

C ===== FIN GRAND BUCLE SOUSSENSAMBLES =====
100 CONTINUE
CALL SHELK (tcorr, 3+2*NAXR, NCACT, 1)
nuls = 0
DO 560 i=1, NCACT
  IF (tcorr(1,i) .EQ. 0) nuls = nuls + 1
560 CONTINUE

C === PREPARER PER HISTOGRAMA ===
ECHEL = 80./tcrut
NETOI = 1. + ECHEL * ABS(tcorr(1,NCACT))
NBLAN = 81 - NETOI

C ===== ECRITURE RESULTATS ELIMINATION =====
WRITE(IMP,891) NV, NV-1, MAXVAR, tcorr(1,NCACT),
- (KETOI,I=1,NETOI), (KBLAN,I=1,NBLAN)

C === ACTUALISATION DE isous; ON CONNAIT LA VAR A ELIMINER: MAXVAR
DO 600 i = 1, NCACT
  isous(MAXVAR,i) = 0
  isous(NCACT+1,i) = isous(NCACT+1,i) + 1
600 CONTINUE
isous(NCACT+1, MAXVAR) = NCACT
C === FIN GRAND-GRAND BUCLE REPETITIONS
50 CONTINUE

C ===== NOU BUCLE DE REMONTAR VARIABLES =====
DO 690 i = 1, NCACT
  isous(i,i) = 1
  IF (isous(NCACT+1,i) .GE. NCACT) THEN
    isous(NCACT+1,i) = NCACT - NFI
    ELSE
      isous(NCACT+1,i) = NCACT
  ENDIF
690 CONTINUE

```

```

C ===== ECRITURE AVANT BUCLE INCORPORATION =====
WRITE (IMP,3000)
WRITE (IMP,892) NCACT
C ===== BIS GRAND-GRAND BUCLE REPETITIONS =====
DO 700 NV = NFI, NCACT
  TCORMAX = -RMAX
  MAXVAR = 0
  DO 701 i = 1, NCACT
    tcorr(1,i) = 0.
    tcorr(2,i) = i
    tcorr(3,i) = 0.
    DO 702 j = 4, 2*NAXR+3
      tcorr(j,i) = 0.
702 CONTINUE
701 CONTINUE

C ===== BIS GRAND BUCLE SOUSENSEMBLES =====
NSUP = NCACT - NV
DO 710 ICAS = 1, NCACT
  IF (isous(NCACT+1,ICAS) .GT. NSUP) GOTO 710

C ===== CHERCHER LE SOUSSENSAMBLE DANS NDIC ET NDON =====
NVARNO = isous(NCACT+1, ICAS)
CALL SOUSDI (NDIC, NSAV, isous(1,ICAS), NCACT, nvsta, 3,
&          NVARNO, 0, nbmod, jorva, NVTOT, jorin, NITOT,
kidm,
&          kvar, kmod, NXMOD, ktitr)
CALL SOUSDO (NDON, NBAND, isous(1,ICAS), NCACT, nvsta, 3,
&          NVARNO, var, NVTOT, kidi, NIDI, ktitr)
C ===== CENTRAGE ET REDUCTION =====
CALL ONDIC (ktitr,NVTOT,NXMODP,NXMACP,nvsta,NMTOT,NMACTP,
&          NITOTP, NIACTP, PITOTP, PIACPT, LANAL, 1, 0,
LERRE,
&          NSAV)
CALL ONDON (ktitr,NVTOT,NIDI,NITOT,TEST,1,0,LERRE,NBAND)
NQEXAP = NVTOT - 1
NQTOTP = nvsta(1) + nvsta(2)
NCONTP = nvsta(3) + nvsta(4)
NCACTP = nvsta(3)
NFTOTP = nvsta(5) + nvsta(6)
ALLOCATE (klib(15,NCONTP), STAT = LERRE)
IF (LERRE.NE. 0) THEN
  write(FICLOG,9000)
  RETURN
ENDIF
READ(NSAV) (nb, l=1, NVTOT+NQTOTP), (jorva(j), j=1, NCONTP)
CALL STABA (evar, ptot, cmoy, cvar, vmin, vmax, ctn, jorva,
&          NCONTP, var, NQEXAP, kidv, NCACTP, NQTOTP,
NIACTP,
&          PIACPT, TEST, PNUL, LERRE, NBAND, NSAV, 0,
NCACT,
&          klib, LEXCE)
DEALLOCATE (klib)
IF(LERRE.NE. 0) RETURN
NISUP = NITOT - NIACT
NISUPP = NITOTP - NIACTP
CALL COREL (ptot, cmoy, cvar, ctn, NCONTP, var, NQEXAP, corr,
&          vmoy, NCACTP, kidi, NIDI, NIACTP, NISUP, PIACPT,
PNUL,
&          TEST, LECORR, NBAND, NSAV, 0)
C ===== DIAGONALISATION =====
CALL VPROP (corr, valpr, ctn, NCACTP, NCACTP, TRAC1, TRAC2, LERRE)
C ===== CALCUL DES CORDONEES DES INDIVIDUS =====
REWIND NBAND
CALL CORIN2(NAXR, NCONTP, NCACTP, NQEXAP, NIDI, NITOTP, valpr,
&          kidi, var, cind, corr, ctn, DOMAX, NSAV, NBAND)
C == CORRELATION AXES: INITIALS(axes01),
axes_sousensamble(axes02) =
REWIND NBAND
CALL OFICH(3, NGUS)
CALL CORAXE (corr, ptot, cind, vmin, cvar, vmax,
&          NAXR, NIACTP, PNUL, TEST, NBAND, NGUS)

C === VERIFIER LE CAS MOINS BON POUR SAVOIR LA VAR A ELIMINER ===
716 CALL CALCRI (valpr, NAXR, corr, tcorr, 3+2*NAXR, NCACT,
ICAS)
IF(tcorr(1,ICAS) .GT. TCORMAX) THEN
  TCORMAX = tcorr(1,ICAS)
  MAXVAR = ICAS
ENDIF
C ===== FIN GRAND BUCLE SOUSSENSAMBLES =====
710 CONTINUE

C === ORDENAR ELS tcorr ===
CALL SHELK (tcorr, 3+2*NAXR, NCACT, 1)
nuls = 0
DO 720 i = 1, NCACT
  IF(tcorr(1,i) .EQ. 0) nuls = nuls + 1
720 CONTINUE

C === PREPARER PER HISTOGRAMA ===

```



```

ECHEL = 80./tcrit
NETOI = 1. + ECHEL * ABS(tcrr(1, NCACT))
NBLAN = 81 - NETOI

C ===== ECRITURE RESULTATS INCORPORATION =====
WRITE(IMP, 891) NV-1, NV, MAXVAR, tcrr(1, NCACT),
- (KETOI, I=1, NETOI), (KBLAN, I=1, NBLAN)

C ACTUALISATION DE isous; ON CONNAIT LA VAR A ELIMINER: MAXVAR =
DO 740 i = 1, NCACT
  isous(MAXVAR, i) = 1
  isous(NCACT+1, i) = isous(NCACT+1, i) - 1
740 CONTINUE
isous(NCACT+1, MAXVAR) = NCACT

C === FIN GRAND-GRAND BUCLE REPETITIONS
700 CONTINUE
RETURN

C ===== FORMATS =====
800 FORMAT (/, ' RECONSTRUCTION DES ', I6, ' PREMIERS AXES', /)
801 FORMAT (/, ' CRITERE INITIAUX AVEC ', I6, ' VARIABLES:
', F10.3, /,
  & ' (somme des correlations carrees des premiers axes',
  & ' ponderees par les valeurs propres)', /)
802 FORMAT (/, ' VALEURS PROPRES DES ', I6, ' PREMIERS AXES: ', /)
803 FORMAT (/, ' LISTE DES ', I6, ' VARIABLES INITIAUX:')
C 801 FORMAT (
C 802 FORMAT (38H ***** INITIALMENT LES CRITERES SONT, F9.2, I8)
C 803 FORMAT (31H et les valeurs propres:)
804 FORMAT (23H *** SOUSENSEMBLES AVEC, I4, 17H VARIABLES ENTRE:)
805 FORMAT (23H *** SOUSENSEMBLES AVEC, I4, 22H VARIABLES PARTINT
DE:)
806 FORMAT (32H *** MEILLEUR SOUSENSEMBLES AVEC, I4, 10H
VARIABLES)
807 FORMAT (
  & 'X ', /, ' ', 'CORRELATIONS ENTRE ELS EIXOS ORIGINALS I ELS
',
  & ' EIXOS RECONSTRUITS:', /,
  & 'X ')
810 FORMAT (50F4.0)
850 FORMAT (15H sans la var , F5.0, F6.2)
851 FORMAT (22H Corr entre eixos, I4, 4H : , F6.4, 6H vap:
, F6.4)
855 FORMAT (15H avec la var , F5.0, F6.2)
860 FORMAT (24H On va , limiter: , F4.0, 9H on
aurr., F9.2, F9.2)
861 FORMAT (24H Les criteres serant: F9.2, F9.2)
865 FORMAT (24H On va ajouter : , F4.0, 9H on
aurr., F9.2, F9.2)
890 FORMAT (/, ' ELIMINATION DES VARIABLES PAS A PAS JUSQUA
', I6, /,
  & ' Step ', ' Var. eliminee', ' Critere ')
891 FORMAT (I4, ' ->', I3, ' VARS.', I10, F10.3, I4, 81A1)
892 FORMAT (/, ' INCORPORATION DES VARIABLES PAS A PAS JUSQUA
', I6, /,
  & ' Step ', ' Var. ajoutee', ' Critere ')
900 FORMAT (I40/1H, 40(1H-)/1H, 40(1H-)/24H SOUSSENSAMBLE
NOMBRE ,
- I3/ 1X, 40(1H-)/ 1X, 40(1H-))
3000 FORMAT (1H1)
4000 FORMAT (1H0/35H Tableau du correlations entre axes)
9000 FORMAT (
FE & 'X ', /, ' ', 'ERREUR (GCOPR-900)', /,
FE & ' ', 'MEMOIRE INSUFFISANTE POUR ALLOUER LE TABLEAU KLIB
', /,
FE & ' ', 'SUPPRIMEZ L'OPTION FICHIER TABLEUR (LEXCE=1)', /,
FE & 'X ')
FE
END

SUBROUTINE SOUSDI(NDICE, NDICS, isous, NSOUS, nvsta, ITYPE,
& NVARNO, NMODNO, nbmod, jorva, NVTOT, jorin,
NITOT,
& kidm, kvar, kmod, NXMOD, ktitre)

C ===== DECLARATIONS EXPLICITES ET DIMENSIONNEMENT =====
INTEGER NDICZ, NDONZ, NDICA, NDONA, NDICB, NDONB, NDIC,
& NDON, NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR
INTEGER LEC, ILLEC, IMP, NXCPA, NXLPA, NLIMP, NDIC1,
& NDON1, NGUS1, NGUS2
INTEGER NDICE, NDICS, NSOUS, ITYPE, NVARNO, NMODNO, NVTOT,
& NITOT, NXMOD, NMOD, NIACT
INTEGER nvsta(8), isous(NSOUS), nbmod(NVTOT), jorva(NVTOT),
& jorin(NITOT)
INTEGER NVTOTE, NXMAC, NMTOTE, NMACTE, LANAL, LERRE, NVSTAE,
& NVSTAS, NVTOTS, NMTOTS, NMACTS, N1, N2, ITETEE,
ITETES
INTEGER i, j, l

CHARACTER ktitre(20), kidm(NXMOD), kvar(15), kmod(5, NXMOD)
REAL PITOT, PIACT

C ===== COMMON /NUFIC/ NDICZ, NDONZ, NDICA, NDONA, NDICB, NDONB, NDIC,
& NDON, NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR
COMMON /ENSOR/ LEC, ILLEC, IMP, NXCPA, NXLPA, NLIMP
COMMON /NGREF/ NDIC1, NDON1, NGUS1, NGUS2

C ===== PROCEDURE =====
CALL ONDIC (ktitre, NVTOTE, NXMOD, NXMAC, nvsta, NMTOTE, NMACTE,
& NITOT, NIACT, PITOT, PIACT, LANAL, 1,
0, LERRE, NDICE)
NVSTAE = nvsta(ITYPE)
NVSTAS = nvsta(ITYPE) - NVARNO
nvsta(ITYPE) = nvsta(ITYPE) - NVARNO
NVTOTS = NVTOTE - NVARNO
NMTOTS = NMTOTE - NMODNO
NMACTS = NMACTE - NMODNO
CALL ONDIC (ktitre, NVTOTS, NXMOD, NXMAC, nvsta, NMTOTS, NMACTS,
& NITOT, NIACT, PITOT, PIACT, LANAL, 0,
0, LERRE, NDICS)

N1 = 0
DO 100 i = 1, ITYPE-1
  N1 = N1 + nvsta(i)
100 CONTINUE
N2 = 0
DO 110 i = ITYPE+1, 6
  N2 = N2 + nvsta(i)
110 CONTINUE
READ (NDICE) (nbmod(1), l=1, NVTOTE), (jorva(j), j=1, NVTOTE-1),
& (jorin(1), l=1, NITOT)
j = 0
DO 200 i = 1, NVSTAE
  IF (isous(i) .EQ. 1) THEN
    j=j+1
    jorva(N1+j) = jorva(N1+i)
  ENDIF
200 CONTINUE
ITETEE = N1 + NVSTAE
ITETES = N1 + NVSTAS
DO 210 i = 1, N2
  nbmod(ITETES+i) = nbmod(ITETEE+i)
  jorva(ITETES+i) = jorva(ITETEE+i)
210 CONTINUE
nbmod(NVTOTE-NVARNO) = nbmod(NVTOTE)
WRITE(NDICS) (nbmod(1), l=1, NVTOTS), (jorva(j), j=1, NVTOTS-1),
& (jorin(1), l=1, NITOT)
DO 300 i = 1, N1
  READ(NDICE) NMOD, (kidm(1), l=1, NMOD), (kvar(1), l=1, 15),
  & ((kmod(1,j), l=1, 5), j=1, NMOD)
  WRITE(NDICS) NMOD, (kidm(1), l=1, NMOD), (kvar(1), l=1, 15),
  & ((kmod(1,j), l=1, 5), j=1, NMOD)
300 CONTINUE
DO 400 i = 1, NVSTAE
  IF (isous(i) .EQ. 1) THEN
    READ(NDICE) NMOD, (kidm(1), l=1, NMOD), (kvar(1),
l=1, 15),
  & ((kmod(1,j), l=1, 5), j=1, NMOD)
    WRITE(NDICS) NMOD, (kidm(1), l=1, NMOD), (kvar(1), l=1, 15),
  & ((kmod(1,j), l=1, 5), j=1, NMOD)
  ELSE
    READ(NDICE) NMOD, (kidm(1), l=1, NMOD), (kvar(1),
l=1, 15),
  & ((kmod(1,j), l=1, 5), j=1, NMOD)
  ENDIF
400 CONTINUE
DO 500 i = 1, N2
  READ(NDICE) NMOD, (kidm(1), l=1, NMOD), (kvar(1), l=1, 15),
  & ((kmod(1,j), l=1, 5), j=1, NMOD)
  WRITE(NDICS) NMOD, (kidm(1), l=1, NMOD), (kvar(1), l=1, 15),
  & ((kmod(1,j), l=1, 5), j=1, NMOD)
500 CONTINUE
RETURN
END

SUBROUTINE SOUSDO (NDONE, NDONS, isous, NSOUS, nvsta, ITYPE,
& NVARNO, var, NVTOT, kidi, NIDI, ktitre)

C ===== DECLARATIONS EXPLICITES ET DIMENSIONNEMENT =====
INTEGER NDICZ, NDONZ, NDICA, NDONA, NDICB, NDONB, NDIC,
& NDON, NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR
INTEGER LEC, ILLEC, IMP, NXCPA, NXLPA, NLIMP, NDIC1, NDON1,
& NGUS1, NGUS2
INTEGER NDONE, NDONS, NSOUS, ITYPE, NVARNO, NVTOT, NIDI
INTEGER isous(NSOUS), nvsta(8), kidi(NIDI)
INTEGER LERRE, NITOT, NVTOTS, N1, N2, ITETEE, ITETES
INTEGER i, ii, j, l

CHARACTER ktitre(20)
REAL var(NVTOT-1)
REAL TEST, POIDS

C ===== COMMON /NUFIC/ NDICZ, NDONZ, NDICA, NDONA, NDICB, NDONB, NDIC,
& NDON, NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR
COMMON /ENSOR/ LEC, ILLEC, IMP, NXCPA, NXLPA, NLIMP
COMMON /NGREF/ NDIC1, NDON1, NGUS1, NGUS2

C ===== PROCEDURE =====
CALL ONDON (ktitre, NVTOT, NIDI, NITOT, TEST, 1, 0, LERRE, NDONE)
IF (LERRE .NE. 0) GOTO 500
NVTOTS = NVTOT - NVARNO

```

```

CALL ONDON (ktitre,NVTOTS,NIDI,NITOT,TEST,0,0,LERRE,NDONS)
IF (LERRE.NE. 0) GOTO 500
N1 = 0
DO 100 i = 1, ITYPE-1
  N1 = N1 + nvsta(i)
100 CONTINUE
N2 = 0
DO 110 i = ITYPE+1, 6
  N2 = N2 + nvsta(i)
110 CONTINUE
DO 200 i = 1, NITOT
  READ(NDONE) (var(1), l=1, NVTOTS-1+NVARNO), POIDS,
    & (kidi(1), l=1, NIDI)
  j = 0
DO 250 ii = 1, nvsta(ITYPE)+NVARNO
  IF (iscus(ii).EQ. 1) THEN
    j = j + 1
    var(N1+j) = var(N1+ii)
  ENDIF
250 CONTINUE
ITETEE = N1 + nvsta(ITYPE) + NVARNO
ITETES = N1 + nvsta(ITYPE)
DO 260 ii = 1, N2
  var(ITETES+ii) = var(ITETEE+ii)
260 CONTINUE
WRITE(NDONS) (var(1), l=1, NVTOTS-1), POIDS,
  & (kidi(1), l=1, NIDI)
200 CONTINUE
500 RETURN
END

SUBROUTINE CORAXE (corr, ptot, cind, cind0, cvar, cvar0,
  & NAXE, NIACT, PNUL, TEST, NFCINO, NFCIN)

C ===== DECLARATIONS EXPLICITES ET DIMENSIONNEMENT =====

INTEGER NDICZ, NDONZ, NDICA, DONA, NDICB, DONB, NDIC,
  & NDON, NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR
INTEGER LEC, ILLEC, IMP, NKCPA, NXLPA, NLIMP
INTEGER NAXE, NIACT, NFCINO, NFCIN
INTEGER IACT, IAX, JAX, i, j
REAL PNUL, TEST, POIDS, PAVAN, COEFM, COEFV, PP
REAL corr(NAXE,NAXE), ptot(NAXE), cind(NAXE), cind0(NAXE),
  & cvar0(NAXE), cvar(NAXE)

C =====

COMMON /NUFIC/ NDICZ, NDONZ, NDICA, DONA, NDICB, DONB, NDIC,
  & NDON, NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR
COMMON /ENSOR/ LEC, ILLEC, IMP, NKCPA, NXLPA, NLIMP

C ===== INITIALISATIONS =====

DO 20 i = 1, NAXE
  ptot(i) = 0.
  cvar(i) = 0.
  cvar0(i) = 0.
DO 10 j = 1, NAXE
  10 corr(i,j) = 0
20 CONTINUE

C ===== LECTURE ET CALCULS =====

DO 50 IACT = 1, NIACT
C ===== LECTURE PAR INDIVIDUS =====
  READ(NFCINO) (cind0(i), i = 1, NAXE), POIDS
  READ(NFCIN) (cind(i), i = 1, NAXE)
  IF (POIDS.LE. PNUL) GOTO 50
C ===== CALCUL DES COEFFICIENTS =====
  DO 40 IAX = 1, NAXE
    IF(cind0(IAX).EQ. TEST) GOTO 40
    PAVAN = ptot(IAX)
    ptot(IAX) = PAVAN + POIDS
    COEFM = POIDS / ptot(IAX)
    COEFV = PAVAN * COEFM
    cvar0(IAX) = cvar0(IAX) + COEFV * cind0(IAX) * cind0(IAX)
    cvar(IAX) = cvar(IAX) + COEFV * cind(IAX) * cind(IAX)
  DO 30 JAX = 1, NAXE
    30 corr(IAX,JAX) = corr(IAX,JAX) + COEFV
  *cind0(IAX)*cind(JAX)
  40 CONTINUE
  50 CONTINUE
C ===== CORRELATION ET ECARTS-TYPES =====

DO 60 IAX = 1, NAXE
  DO 55 JAX = 1, NAXE
    PP = SQRT(cvar0(IAX)) * SQRT(cvar(JAX))
    55 corr(IAX,JAX) = corr(IAX,JAX) / PP

```

```

60 CONTINUE
  DO 70 IAX = 1, NAXE
    cvar(IAX) = SQRT(cvar0(IAX) / ptot(IAX))
    cvar(IAX) = SQRT(cvar(IAX) / ptot(IAX))
  70 CONTINUE
END

SUBROUTINE CALCRI(valpr, NAXE, corr, tcorr, NLIG, NCACT, ICAS)

C ===== DECLARATIONS EXPLICITES ET DIMENSIONNEMENT =====

INTEGER NAXE, NLIG, NCACT, ICAS
INTEGER LEC, ILLEC, IMP, NKCPA, NXLPA, NLIMP, i
REAL valpr(NAXE), corr(NAXE,NAXE), tcorr(NLIG,NCACT)

C =====

COMMON /ENSOR/ LEC, ILLEC, IMP, NKCPA, NXLPA, NLIMP

C ===== PROCEDURE =====

DO 100 i = 1, NAXE
  tcorr(1,ICAS) = tcorr(1,ICAS) + valpr(i) * corr(i,i) * corr(i,i)
  tcorr(3,ICAS) = tcorr(3,ICAS) + corr(i,i) * corr(i,i)
  tcorr(3+i,ICAS) = corr(i,i)
  tcorr(3+NAXE+i,ICAS) = valpr(i)
100 CONTINUE
RETURN
END

SUBROUTINE CORINZ (NAXE,NCONT1,NCACT1,NQEXA1,NIDI1,NITOT1,valpr,
  & kidil,var1,cind,ccon, ctn,DOMAX,NFSAV,NFOUT)

C ===== DECLARATIONS EXPLICITES ET DIMENSIONNEMENT =====

INTEGER NAXE, NCONT1, NCACT1, NQEXA1, NIDI1, NITOT1,
NFSAV,
  & NFOUT
INTEGER NDICZ, NDONZ, NDICA, DONA, NDICB, DONB, NDIC, NDON,
  & NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR
INTEGER LEC, ILLEC, IMP, NKCPA, NXLPA, NLIMP
INTEGER NDIC1, DON1, NGUS1, NGUS2
INTEGER IZERO, IND, IAX, IC, i
CHARACTER kidil(NIDI1)
REAL DOMAX, CRD, DORIG, POIDS
REAL valpr(NAXE), var1(NQEXA1), cind(NAXE), ccon(NCACT1,NAXE),
  & ctn(NCONT1)

C =====

COMMON /NUFIC/ NDICZ, NDONZ, NDICA, DONA, NDICB, DONB, NDIC,
  & NDON, NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR
COMMON /ENSOR/ LEC, ILLEC, IMP, NKCPA, NXLPA, NLIMP
COMMON /NGREF/ NDIC1, DON1, NGUS1, NGUS2

C ===== PROCEDURE =====

REWIND NFSAV
DOMAX = 0.
IZERO = 0.

C ===== CORDONEES FACTORIALLES =====

DO 70 IND = 1, NITOT1
  READ(NFSAV) (var1(i), i=1, NQEXA1), POIDS,
    & (kidil(i), i=1, NIDI1)
  DO 50 IAX = 1, NAXE
    CRD = 0.
    IF (valpr(IAX).EQ. 0.) GOTO 50
    DO 40 IC = 1, NCACT1
      40 CRD = CRD + ctn(IC) * ccon(IC,IAX)
    50 cind(IAX) = crd

C ===== DISTANCE A L'ORIGINE =====

DORIG = 0.
DO 60 IC = 1, NCACT1
  60 DORIG = DORIG + ctn(IC) * ctn(IC)
DOMAX = AMAX1(DOMAX,DORIG)

C ===== SAUVEGARE SUR NFOUT =====

WRITE(NFOUT) (cind(i), i=1, NAXE), POIDS, DORIG,
  & (kidil(i), i = 1, NIDI1), IZERO
  70 CONTINUE
RETURN
END

```

Les subrutines COREL i STABA, encarregades de calcular la matriu de correlacions i les estadístiques descriptives de les variables estan estretes de la llibreria proporcionada per SPAD, dins del procediment COPRI.

Annex2. Programa GREF1 complet

```
C*****
C***** GREF1 *****
C*****
```

```
SUBROUTINE DUMM1
.NXMQ, NXKQ, NXRQ, NXDQ, LISTF, LRESM, LERRE, LEPAR,
&
& KFICH, NFICH, LEXCE)

C***** DECLARATIONS EXPLICITES ET DIMENSIONNEMENT *****
```

```
INTEGER NXMQ, NXKQ, NXRQ, NXDQ, LISTF, LRESM, LERRE
INTEGER LEXCE, LEPAR, NFICH
INTEGER LEC, ILEC, IMP, NXCPA, NXLPA, NLIMP
INTEGER NDICZ, NDONZ, NDICA, NDONA, NDICB, NDONB, NDIC, NDON,
& NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR
& INTEGER MDICZ, MDONZ, MGUSZ, MGUSG, NMACZ, NMACA,
& NTRA, NTRA2, NBASE
INTEGER NCOEF, NCOEFB
INTEGER FICLOG
INTEGER IERR
INTEGER NDIC1, NDON1, NGUS1
INTEGER LEDIN, LTEST, NID, NCON,
& NCONT1, NQEX, NQEXA1, IQNB1, IQJRV1, IKID11,
& IDMORI, IKIDM, IKVAR, IKMOD, IKARW, IKIDV, NKARW,
& NCACT1, IVALP, IRACVA, IGRAV, ICIND, IRVAR1
& ICTN1, NQTOT1, ICCON, ICCONT, ICORR, IPTOT, ICMOY,
& ICVAR, ICTN, IEVAR, ICVARAX, IVMIN, IVMAX, IVMOY,
& ICMODT, IPMODT,
& IEMODT, IVMOD, IVEPS0, IVEPS1, ICORRV,
& NAXED
INTEGER IQNMCUM, IQJRV
INTEGER NVTOT1, NITOT1, NIACT1
INTEGER NXMOD1, NXMAC1, NMTOT1, NMACT1, NIDI1
INTEGER NQEXA, NQTOT, NQACT, NCONT, NCACT, NPTOT
INTEGER NIDI, NAXE, LORIG, NIDV
INTEGER NVTOT, NXMOD, NXMAC, NMTOT, NMACT, NITOT, NIACT, LANAL
INTEGER NQACT1, NPTOT1
INTEGER NVSTA1(8), MPARM(1), NLTYPI(12)
INTEGER NLTYPI(12), I
CHARACTER*4 KTITRE(20)
CHARACTER*4 KETA1, KETA2, KANAL
CHARACTER*4 KTITR(20)
CHARACTER*256 KFICH(NFICH)
```

```
REAL PITOT, PIACT, TEST
REAL PITOT1, PIACT1, TEST1
REAL VBIID(1)
```

```
C ===== TABLEAUX EN ALLOCATION DYNAMIQUE =====
```

```
INCLUDE 'DINTAB.INC'
COMMON /ENSOR/ LEC, ILEC, IMP, NXCPA, NXLPA, NLIMP
COMMON /NFIC/ NDICZ, NDONZ, NDICA, NDONA, NDICB, NDONB,
& NDIC, NDON,
& NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR
COMMON /NOFIC/ MDICZ, MDONZ, MGUSZ, MGUSG, NMACZ, NMACA,
& NTRA, NTRA2, NBASE
COMMON /ILOG/ FICLOG
COMMON /NDISC/ NCOEF, NCOEFB
COMMON /NGREF/ NDIC1, NDON1, NGUS1
DATA KETA1 /'DUMM' /, KETA2 /'1' /
```

```
C ===== ALLOCATION POUR LECTURE DES PARAMETRES =====
```

```
ALLOCATE (MQ(NXMQ), STAT=LERRE)
IF (LERRE .NE. 0) GOTO 901
ALLOCATE (KQ(NXKQ), STAT=LERRE)
IF (LERRE .NE. 0) GOTO 901
ALLOCATE (RQ(NXRQ), STAT=LERRE)
IF (LERRE .NE. 0) GOTO 901
ALLOCATE (DQ(NXDQ), STAT=LERRE)
IF (LERRE .NE. 0) GOTO 901
DQ(1) = 0.0
CALL DEBET (KETA1, KETA2, KTITR)
NKARW = 35
```

```
C===== OUVERTURE DU FICHIER NDIC
```

```
CALL ONDIC (KQ, NVTOT, NXMOD, NXMAC, MQ, NMTOT, NMACT,
- NITOT, NIACT, PITOT, PIACT, LANAL, 1, LISTF, LERRE, NDIC)
IF (LERRE .NE. 0) GOTO 40
NQEXA = NVTOT - 1
NQTOT = MQ(1) + MQ(2)
NQACT = MQ(1)
NCONT = MQ(3) + MQ(4)
NCACT = MQ(3)
NPTOT = MQ(5) + MQ(6)
```

```
C===== OUVERTURE DU FICHIER NDON
```

```
CALL ONDON (KQ, NVTOT, NIDI, NITOT, TEST, 1, 1, LERRE,
& NDON)
IF (LERRE .NE. 0) GOTO 40
```

```
C===== OUVERTURE DU FICHIER NGUS
```

```
CALL ONGUS (KTITR, NAXE, LORIG, NIDI, NIDV, NLTYPI, 1,
- LISTF, LERRE, NGUS, KANAL)
IF (LERRE .NE. 0) GOTO 40
```

```
C===== OUVERTURE DU FICHIER NDIC1 =====
CALL ONDIC
(KQ, NVTOT1, NXMOD1, NXMAC1, NVSTA1, NMTOT1, NMACT1, NITOT1,
- NIACT1, PITOT1, PIACT1, LANAL, 1, LISTF,
LERRE, NDIC1)
IF (LERRE .NE. 0) GOTO 40
NQEXA1 = NVTOT1 - 1
NQTOT1 = NVSTA1(1) + NVSTA1(2)
NQACT1 = NVSTA1(1)
NCONT1 = NVSTA1(3) + NVSTA1(4)
NCACT1 = NVSTA1(3)
NPTOT1 = NVSTA1(5) + NVSTA1(6)
```

```
C===== OUVERTURE DU FICHIER NDON1 =====
CALL ONDON (KQ, NVTOT1, NIDI1, NITOT1, TEST1, 1, 1,
LERRE, NDON1)
IF (LERRE .NE. 0) GOTO 40
```

```
C===== OUVERTURE DU FICHIER NGUS1 EN ECRITURE =====
```

```
DO 10 I=1, 12
10 NLTYPI(I) = 0
NLTYPI(1) = NIACT1
NLTYPI(2) = NITOT1 - NIACT1
NLTYPI(3) = NMACT1
NLTYPI(4) = NMTOT1 - NMACT1
NLTYPI(5) = NVSTA1(3)
NLTYPI(6) = NVSTA1(4)
NLTYPI(7) = NVSTA1(5)
NLTYPI(8) = NVSTA1(6)
NLTYPI(11) = NLTYPI(11)
CALL ONGUS (KTITRE, NAXE, LORIG, NIDI1, 6, NLTYPI, 0,
- LISTF, LERRE, NGUS1, '2')
IF (LERRE .NE. 0) GOTO 40
```

```
C===== LECTURE DE PARAMETRES =====
```

```
MPARM(1) = NAXE
CALL PARAM (MQ, NXMQ, KQ, NXKQ, RQ, NXRQ, MPARM, 1,
& VBIID, 1, 0, NXDQ, KETA1, LEPAR, LRESM, LERRE)
IF (LERRE .NE. 0) GOTO 40
LEDIN = MQ(1)
NAXED = MINO(MQ(2), NAXE)
LTEST = MQ(3)
```

```
C===== RESERVATION POUR GRACP =====
```

```
NID = MAX0 (NIDI, NIDI1)
NCON = MAX0 (NCONT, NCONT1)
NQEX = MAX0 (NQEXA, NQEXA1)
```

```
C == RESERVATION DE ENTIERS
```

```
IQNB1 = 1
IQJRV1 = IQNB1 + NVTOT1
```

```
C == RESERVATION POUR CORMO2
```

```
IQNMCUM = IQJRV1 + NVTOT1
IQJRV = IQNMCUM + NVTOT1
NMQ = IQJRV + NVTOT
```

```
C == RESERVATION DE CARACTERS
```

```
IKID11 = 1
```

```
C == RESERVATION POUR CORMO2
```

```
IKIDM = IKID11 + NID
IKVAR = IKIDM + NXMOD1
IKMOD = IKVAR + 15
```

```
C == RESERVATION POUR LIRE LEPAR
```

```
IKARW = IKMOD + NXMOD1 * 5
IKIDV = IKARW + NKARW
NKQ = IKIDV + NCACT1
```

```
C == RESERVATION DE REELS
```

```
IVALP = 1
IRACVA = IVALP + NAXE
```

```
C == RESERVATION POUR cntre gravit,e
```

```
IGRAV = IRACVA + NAXE
```

```
C == RESERVATION POUR LIRE NDON1 ET UTILISER ... CORIN2
```

```
ICIND = IGRAV + NAXE
IRVAR1 = ICIND + NAXE
```

```
C == RESERVATION DE SUBPART DE VAR1 POUR GRACP (CTN)
```

```
ICTN = IRVAR1 + NQTOT
ICTN1 = IRVAR1 + NQTOT1
```

```
C == RESERVATION POUR LIRE NGUS ET UTILISER ... CORIN2
```

```
ICCON = IRVAR1 + NQEX
ICCONT = ICCON + NCACT1 * NAXE
ICORR = ICCONT + NCONT1 * NAXE
IPTOT = ICORR + NCACT1 * NCACT1
ICMOY = IPTOT + NCON
ICVAR = ICMOY + NCON
IEVAR = ICVAR + NCON
```

```
C == RESEVATION POUR ETY(axes1) staba apres corin2
```

```
ICVARAX = IEVAR + NCON
```

```
C == RESEVATION POUR STABA ,COREL
```

```
IVMIN = ICVARAX + NAXE
IVMAX = IVMIN + NCON
IVMOY = IVMAX + NCON
```

```
C == RESERVATION POUR CORCI2
```

```
ICMODT = IVMOY + NCACT1
IPMODT = ICMODT + NMTOT1 * NAXE
```

```
C == RESERVATION POUR CORMO2
```

```
IDMORI = IPMODT + NMTOT1
IEMODT = IDMORI + NMTOT1
IVMOD = IEMODT + NMTOT1
```

```
C == RESERVATION POUR CORRELATIONS VEPS
```

```
IVEPS0 = IVMOD + NMTOT1 * NCACT1
IVEPS1 = IVEPS0 + NCACT1 * NAXE
```

```
ICORRV = IVEPS1 + NCACT1 * NAXE
```

```
NRQ = ICORRV + NAXE * NAXE
```

```

NDQ = 1

C ===== DEALLOCATION APRES LECTURE DES PARAMETRES =====
DEALLOCATE (MQ, STAT=IERR)
DEALLOCATE (KQ, STAT=IERR)
DEALLOCATE (RQ, STAT=IERR)
DEALLOCATE (DQ, STAT=IERR)

C ===== ALLOCATION DYNAMIQUE DES TABLEAUX POUR EXECUTION =====
ALLOCATE (MQ(NMQ), STAT=LERRE)
IF (LERRE.NE. 0) GOTO 902
ALLOCATE (KQ(NKQ), STAT=LERRE)
IF (LERRE.NE. 0) GOTO 902
ALLOCATE (RQ(NRQ), STAT=LERRE)
IF (LERRE.NE. 0) GOTO 902
ALLOCATE (DQ(NDQ), STAT=LERRE)
IF (LERRE.NE. 0) GOTO 902

CALL GRACP (LORIG, TEST, TEST1, NKARW, LEPAR, LEDIN, NAXED,
- NVTOT1, NQEXA1, NIACT1, NCONT1, NCACT1, PIACT1, NITOT1,
- NMTOT1, NXMOD1, NAXE, NID11, NLTY, NVSTAL, NQEXA, PIACT,
- MQ(IQNB1), MQ(IQJRV1), MQ(IQNMCM), MQ(IQJRV),
- KQ(IKID1), KQ(IKIDM), KQ(IKVAR), KQ(IKMOD), KQ(IKARW),
- KQ(IKIDV), KQ(IKIDF), KQ(IKACVA), KQ(IKRAV),
- RQ(ICIND), RQ(IRVAR1), RQ(ICCON), RQ(ICCONT),
- RQ(ICORR), RQ(IPTOT), RQ(ICMOY), RQ(ICVAR),
- RQ(IEVAR), RQ(ICTN), RQ(ICTN1), RQ(IVMIN),
- RQ(IVMAX), RQ(IVMOY), RQ(ICVARAX), RQ(ICMODT), RQ(IPMODT),
- RQ(IDMORI), RQ(IEMODT), RQ(IVMOD),
- RQ(IPEPS0), RQ(IPEPS1), RQ(ICORRV),
- NMTOT, NITOT, NIACT, NCONT, NQOTOT, NID, NCON, NQEX,
- NVTOT, LTEST, LEXCE)

GOTO 40

C ===== FIN DE LA PROCEDURE =====
40 CALL FINET (KETAL, KETA2, KTITR)
GOTO 999

C ===== ERREUR D'ALLOCATION =====
901 CALL MALLO (KETAL, KETA2, NMQ, NKQ, NRQ, NXDQ)
GOTO 40
902 CALL MALLO (KETAL, KETA2, NMQ, NKQ, NRQ, NDQ)
GOTO 40

C == DEALLOCATION DES TABLEAUX APRES EXECUTION (SI LERRE = 0) ==
999 IF (LERRE.EQ. 0) THEN
DEALLOCATE (MQ, STAT=IERR)
DEALLOCATE (KQ, STAT=IERR)
DEALLOCATE (RQ, STAT=IERR)
DEALLOCATE (DQ, STAT=IERR)
ENDIF

C ===== FORMATS =====
RETURN
END

SUBROUTINE GRACP (LORIG, TEST, TEST1, NKARW, LEPAR, LEDIN,
- NAXED,
- NVTOT1, NQEXA1, NIACT1, NCONT1, NCACT1, PIACT1, NITOT1,
- NMTOT1, NXMOD1, NAXE, NID11, NLTY, NVSTAL, NQEXA,
- PIACT,
- NBMOD1, JORVA1, NMCUM, JORVA,
- KID11, KIDM1, KVAR1, KMOD1, KARW, KIDV,
- VALPR, RACVA, CCGRAV,
- CIND, VAR1, CCON, CCONT,
- CORR, PTOT, CMOY, CVAR,
- EVAR, CTN, CTN1, VMIN,
- VMAX, VMOY, CVARAX, CMODT, PMODT,
- DMORI, EMODT, VMOD,
- veps0, veps1, corrv,
- NMTOT, NITOT, NIACT, NCONT, NQOTOT, NID, NCON, NQEX, NVTOT, LTEST,
- LEXCE)

INTEGER NCON, NCACT1, NCONT1, NAXE, NCONT, NMTOT1,
NXMOD1,
- NVTOT1, NQEXA, NID, NKARW, NQEX, NXFOR, NPAS
INTEGER I, J, LCCORR, LORIG, LEXCE

INTEGER NDICZ, NDONZ, NDICA, NDONA, NDICB, NDONB,
NDON,
- NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR, FICLOG,
- LEC, ILLEC, IMP, NXCPA, NXLPA, NLIMP,
- NDIC1, NDON1, NGUS1, NGUS2, mdonz, mgusz, mgusg,
- nmacz, nmaca, ntra, ntra2, nbase
INTEGER NINDT, IFOR1, NISUP1, NITOT1, NIACT1, NQOTOT1,
- NCACT, IND, L, LTEST, NPASG, NITOT, NMTOT, NIACT, LERRE,
NQEXA1,
- NID11, IAX, NCILL1, IFOR2, IFOR3
INTEGER NVTOT, NQOTOT, NB, LEPAR, NAXED, LEDIN, LCCOS2,
- IC
INTEGER NLTY(12), NVSTAL(8),
- NBMOD1(NVTOT1), JORVA1(NVTOT1), NMCUM(NVTOT1), JORVA(NQEXA)
CHARACTER*4 KFACT
CHARACTER*4 KID11(NID), KARW(NKARW), KIDM1(NXMOD1),
- KVAR1(15), KMOD1(5, NXMOD1), KIDV(NCACT1)
CHARACTER*4 KLIB(:,:)

```

```

REAL ZERO, RMAX
REAL TRACE, P, D, PNUL1, PIACT1, DCDGO, POIDS,
- TEST, SVAL, DOMAX, TEST1, PCGRAV, TRAC1, TRAC2,
- PIACT, VANEG, PNUL, RMIN
REAL VALPR(NAXE), RACVA(NAXE), CIND(NAXE), VAR1(NQEX)
- CCON(NCACT1, NAXE),
- CORR(NCACT1, NCACT1), PTOT(NCON),
- CCONT(NCONT1, NAXE), CMOY(NCON), CVAR(NCON), EVAR(NCON),
- VMOY(NCACT1),
CTN1(NCONT1), VMIN(NCON), VMAX(NCON),
- CVARAX(NAXE), CTN(NCONT), CCGRAV(NAXE),
- CMODT(NMTOT1, NAXE),
- PMODT(NMTOT1), EMODT(NMTOT1),
- DMORI(NMTOT1),
- VMOD(NMTOT1, NCACT1),
- veps0(naxe, ncact1), veps1(naxe, ncact1),
- corrv(naxe, naxe)

COMMON /SEUIL/ ZERO, RMIN, RMAX
COMMON /NUFIC/ NDICZ, NDONZ, NDICA, NDONA, NDICB, NDONB,
- NDIC, NDON, NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR
COMMON /ENSOR/ LEC, ILLEC, IMP, NXCPA, NXLPA, NLIMP
COMMON /NGREF/ NDIC1, NDON1, NGUS1
COMMON /NGUS2/ NGUS2
COMMON /ILOG/ FICLOG
common /nofic/ mdicz, mdonz, mgusz, mgusg, nmacz, nmaca,
- ntra, ntra2, nbase

DATA NXFOR /20/, IFOR1 /7/, IFOR2 /9/, IFOR3 /10/
KFACT = 'F7.2'
ALLOCATE (KLIB(15, NCACT), STAT = LERRE)
if (lerre.ne. 0) then
write (ficlog, 9000)
return
endif

C ===== LECTURE SUR NGUS =====
CALL OFICH(2, NGUS)
READ (NGUS) (VALPR(I), I=1, NAXE), TRACE

C ===== ECRITURE SUR NGUS1 =====
WRITE(NGUS1) (VALPR(I), I=1, NAXE), TRACE
WRITE(NGUS2) (VALPR(I), I=1, NAXE), TRACE

C ===== LECTURE DES VARIABLES CONTINUES ACTIVES SUR NGUS =====
C ===== POUR UTILISER A CORIN2 =====
NINDT = NLTY(1) + NLTY(2) + NLTY(3) + NLTY(4)
CALL OFICH (3+NINDT, NGUS)
DO 100 IC = 1, NCACT1
READ (NGUS) (CCON(IC, I), I=1, NAXE), P, D
100 CONTINUE

C ===== INITIALISATIONS =====
LCCORR = 2 - LORIG
PNUL = VANEG(PIACT)
PNUL1 = VANEG(PIACT1)
NISUP1 = NITOT1 - NIACT1
NQOTOT1 = NVSTAL(1) + NVSTAL(2)
NCILL1 = NVSTAL(4)

C ===== CENTRAGE (ET REDUCTION) =====
C == LECTURE SUR NDIC1 ==
NCACT=NCACT1
READ (NDIC) (NB, L=1, NVTOT+NQOTOT), (JORVA(J), J=1, NCONT)
CALL STABA (EVAR, PTOT, CMOY, CVAR, VMIN, VMAX, CTN, JORVA,
- NCACT, VAR1, NQEXA, KIDV, NCACT, NQOTOT,
- NIACT, PIACT, TEST, PNUL, LERRE, NDON, NDIC, 1,
- NCACT, KLIB, LEXCE)
IF (LERRE.NE. 0) RETURN
CALL COREL (PTOT, CMOY, CVAR, CTN1, NCACT1, VAR1, NQEXA1, CORR,
- VMOY, NCACT1, KID11, NID11, NIACT1, NISUP1, PIACT1,
- PNUL1, TEST1, LCCORR, NDON1, NSAV)

C ===== CALCUL DES COORDONNEES DES INDIVIDUS =====
C == TRANSFORMATION DES COORDONNEES ==
DO 112 IAX = 1, NAXE
SVAL = SQRT(VALPR(IAX))
DO 112 I = 1, NCACT1
112 CCON(I, IAX) = CCON(I, IAX) / SVAL
CALL COIN2 (NAXE, NCONT1, NCACT1, NQEXA1, NID11,
- NITOT1, VALPR, KID11, VAR1,
- CIND, CCON, CTN1, DOMAX, NSAV, NGUS1)

C == CALCUL CENTRE GRAVIT, E ==
DO 120 IAX=1, NAXE
CCGRAV(IAX)=0
120 CONTINUE
PCGRAV=0
CALL OFICH(3, NGUS1)
DO 121 IND=1, NIACT1
READ (NGUS1) (cind(I), I=1, NAXE), POIDS
DO 122 IAX=1, NAXE
CCGRAV(IAX)=CCGRAV(IAX)+CIND(IAX)*POIDS
122 CONTINUE
PCGRAV=PCGRAV+POIDS
121 CONTINUE
DO 123 IAX=1, NAXE
CCGRAV(IAX) = CCGRAV(IAX) / PCGRAV
123 CONTINUE

C == CALCUL DISTO REEL CENTRE GRAVIT, E ==
CALL DICG2 (NIACT1, NCACT1, PIACT1, VAR1, NQEXA1, CTN1, DCDGO,

```

```

- NSAV,VMOD)

C ===== CENTRAGE COORD.(CENTRAGE AXES):code comme staba et
corel =====
CALL STBCO (EVAR,PTOT,CMOY,CVARAX,VMIN,VMAX,NCON,NAXE,
- CIND,NITOT1,NIAC1,LCORR,PNU1,NGUS1,NBAND)

C ===== CALCUL DES COORDONN,ES DES CONTINUES ACT. ET ILL. =====
C ===== (centrage,comparation corr,calcul et edition) =====

C === CENTRAGE (ET REDUCTION) par rapport ... analyse 1 ===
READ(NDIC1) (NB,L=1,NVTOT1+NQTOT1),
(JORVAL(J),J=1,NCONT1)
CALL OFICH (Z,NDON1)
CALL STABA (EVAR,PTOT,CMOY,CVAR,VMIN,VMAX,CTN1,JORVAL1,
- NCONT1,VARI,NQEXAL,KIDV,NCAC1,NQTOT1,
-
- NIAC1,PIAC1,TEST1,PNU1,LERRE,NDON1,NDIC1,1,
- NCAC1,KLIB,LEXCE)
IF (LERRE.NE.0) RETURN
DEALLOCATE(KLIB)

C == sorties centrage et reduction ->EVAR PTOT CMOY CVAR
CALL COREL (PTOT,CMOY,CVAR,CTN1,NCONT1,VARI,NQEXAL,CORR,
- VMOY,NCAC1,KID1,NID1,NIAC1,NISUP1,PIAC1,
- PNU1,TEST1,LCORR,NDON1,NTRA)

C ->>>donnees centr,es+reduites par rapport ... analyse 1 ->
NTRA

C +++ garder corr sur NTRA +++
NPAS=NIAC1+NISUP1
CALL OFICH (NPAS,NTRA)
DO 127 I=1,NCAC1
DO 127 J=1,NCAC1
127 WRITE(NTRA) CORR(I,J)

C == COMPARATION des matrices des correlations =====
IF (LTEST.EQ.1) THEN
NPASG = 3+NITOT+NMTOT+NCONT+NCAC1
C == NCAC1=NCAC1
C la lecture de jorval ... et, fet avant staba
CALL TESTCO(CORR,VMIN,VMAX,KIDV, NCAC1,
- NIAC1,NIAC1,JORVAL1,
- NGUS,NPASG,NTRA,NPAS)

C +++ recuperer corr du NTRA qui a ete detruite par TESTCO
+++
NPAS=NIAC1+NISUP1
CALL OFICH (NPAS,NTRA)
DO 128 I=1,NCAC1
DO 128 J=1,NCAC1
128 READ(NTRA) CORR(I,J)
+++++++
ENDIF
C endif du if LTEST.EQ.1

C == CALCUL DES CORRELATIONS ENTRE VEPS =====
NPAS=NLTYP(1)+NLTYP(2)+NLTYP(3)+NLTYP(4)+NLTYP(5)+NLTYP(6)
CALL OFICH (3+NPAS,NGUS)
DO 129 I=1,NCAC1
READ(NGUS) (CCON(I,J),J=1,NAXE)
129 CONTINUE
CALL TRASP (CCON,NCAC1,NAXE,VEPS0)

C +++ calculer veps1 avec CALL VPROP(CORR,...)
CALL VPROP (CORR,VALPR,CTN,NCAC1,NCAC1,TRAC1,TRAC2,LERRE)
CALL TRASP (CORR,NCAC1,NAXE,VEPS1)

C +++ comprovations des signes des veps
CALL SIGNE2 (VEPS0,NAXE,VEPS1,NAXE,NCAC1,NAXE,CCON,CCONT)

C +++ correlation comme produit scalaire:cos2
C matrice avec les correlations:ccont
C veps1 -> veps1t=ccon
CALL TRASP (VEPS1,NAXE,NCAC1,CCON)
CALL MULTM (VEPS0,NAXE,NCAC1,CCON,NCAC1,NAXE,NCAC1,CORRV)
WRITE(IMP,3000)
NLIMP=1
C WRITE(IMP,4500)
WRITE(IMP,4000)
NLIMP=NLIMP+4
WRITE(IMP,5000)
do 130 I=1,naxe
write(imp,6000) (CORRV(I,J),J=1,naxe)
130 continue
WRITE(IMP,5500)
C +++ recuperer corr du NTRA qui a ete detruite par VPROP
+++
NPAS=NIAC1+NISUP1
CALL OFICH (NPAS,NTRA)
DO 131 I=1,NCAC1
DO 131 J=1,NCAC1
131 READ(NTRA) CORR(I,J)

C == CALCUL des coordonn,es et EDITION =====
CALL CORCON ( NAXE,NIAC1,NQEXAL,PIAC1,NCONT1,
- CCONT,VARI,CIND,CTN1,CVARAX,
- EVAR,PTOT,CMOY,CVAR,NTRA,NBAND)
READ(LEPAR,1000) (KARW(I),I=1,NXFOR)
KARW(IFOR1) = KFACT
DO 132 I=1,NAXE
RACVA(I) = 0
132 CONTINUE
CALL ECCVA
(RACVA,NAXED,CVAR,NAXE,KARW,NXFOR,NCAC1,NCILL1,
- NQTOT1,LCORR,NDIC1,lexce,lerre)
C ===== EDITIONS DES RESULTATS selon type =====

```

```

C ===== EDITION DES INDIVIDUS =====
READ(LEPAR,1000) (KARW(I),I=1,NXFOR)
KARW(IFOR2) = KFACT
IF (LEDIN.NE.0) THEN
cjan94 posar 0 perque surtin cont abs 0.ho posem a racva
do 134 i=1,naxe
racva(i)=0
134 continue
IF (LCORR.NE.0) THEN
LCOS2=0
CALL ECCIG (CIND,NAXE,racva,NAXED,PIAC1,
- NIAC1,NISUP1,NID1,LEDIN,NGUS1,LCOS2)
ENDIF
IF (LCORR.EQ.0) THEN
LCOS2=0
CALL ECING (CIND,NAXE,racva,NAXED,KARW,NXFOR,
-
- PIAC1,NIAC1,NISUP1,NID1,LEDIN,DOMAX,NGUS1,LCOS2)
ENDIF
DO 135 I=1,NAXE
RACVA(I) = SQRT (VALPR(I))
135 CONTINUE
ENDIF

Coct94 == editar centre grav es independent d'edicio 1 o 2 !
C ===== EDITION CENTRE GRAVIT,E
=====
C == mars94 editar centre grav es independent d'edicio 1 o
2 ??
CALL EDCDG(ccgrav,NAXED,VALPR,racva,PIAC1,PIAC1,DCDGO)

C == CALCUL DES COORDONN,ES DES MODALIT,ES ILLUSTRATIVES
=====
C == LECTURE SUR NDIC1 ==
CALL OFICH (2,NDIC1)
READ(NDIC1) (NBMOD1(L),L=1,NVTOT1),
(JORVAL(J),J=1,NVTOT1)
IF (NMTOT1.NE.0) THEN
NMCUM(1) = 0
DO 138 I = 1,NQEXAL
NMCUM(I+1) = NMCUM(I) + NBMOD1(I)
138 CONTINUE
CALL CORMO2 (NMCUM,NQTOT1,CMODT,PMODT,EMODT,
- NMTOT1,NVTOT1,NQEXAL,NIAC1,NCAC1,
- NAXE,CIND,VARI,PNU1,
- DMORI,
- VMOD,CTN1,NCONT1,NDIC1,NDON1,NGUS1,NBAND,NSAV)

C ===== ET EDITION =====
=====
READ(LEPAR,1000) (KARW(I),I=1,17)
C KARW(IFOR3) = KFACT
CALL EVTES
(CMODT,PMODT,EMODT,DMORI,VALPR,RACVA,NMTOT1,NAXED,
-
- NMCUM,NVTOT1,JORVAL1,NQTOT1,KIDM1,KMOD1,NXMOD1,
-
- KARW,17,KVARI,NIAC1,PIAC1,LORIG,PNU1,NDIC1)
C -
KARW,15,KVARI,NIAC1,PIAC1,LORIG,PNU1,NDIC1)

ENDIF

C ----- FORMATS -----
1000 FORMAT (20A4)
3000 FORMAT (1H1)
4000 FORMAT (1H0/43H CORRELATIONS ENTRE AXES DU NGUS ET
NGUS1/
- 3X,41(1H-))
4500 FORMAT (1H0/38H COSINUS ENTRE AXES DU NGUS ET NGUS1/
- 1X,41(1H-))
5000 FORMAT (4X,1H|,15H Axes du NGUS1/
- 4H---,1H+,60(1H-))
5500 FORMAT (4H---,1H+,60(1H-)/
- 4X,1H|)
6000 FORMAT (4X,'|',15(F7.2))
9000 FORMAT (
FE & 'X',/,',', 'ERREUR (GCOPR-900)',/,
FE & ' ', 'MEMOIRE INSUFFISANTE POUR ALLOUER LE TABLEAU KLIB
',/,
FE & ' ', 'SUPPRIMEZ L'OPTION FICHIER TABLEUR (LEXCE=1)',/,
FE & 'X ')
FE
C -----
END
SUBROUTINE COIN2 (NAXE,NCONT1,NCAC1,NQEXAL,NID1,
- NITOT1,VALPR,KID1,VAR1,
- CIND,CCON,CTN,DOMAX,NFSAV,NFOUT)

INTEGER NAXE, NCONT1, NCAC1, NQEXAL, NID1, NITOT1
INTEGER LEC, ILLEC, IMP, NXCPA, NXLPA, NLIMP, NDIC2,
- NDONZ, NDICA, NDONA, NDICB, NDONB, NDIC, NDON, NGUS,
- NGRI, NGRO, NSAV, NBAND, NBFOR, NDIC1, NDON1, NGUS1
INTEGER NFSAV, IZERO, IND, I, IAX, IC, NFOUT
CHARACTER*4 kidil(nidil)
REAL POIDS, DOMAX, DORIG, CRD
REAL VALPR(NAXE), VAR1(NQEXAL),
- CIND(NAXE),CCON(NCAC1,NAXE), CTN(NCONT1)

```

```

COMMON /ENSOR/ LEC,ILLEC, IMP,NXCPA,NXLPA,NLIMP
COMMON /NUFIC/ NDICZ,NDONZ, NDICA,NDONA, NDICB,NDONB,
NDIC,NDON,
- NGUS,NGRI,NGRO, NSAV,NBAND,NBFOR
common /ngref/ ndic1,ndon1,ngus1

REWIND NFSAV
DOMAX = 0.
IZERO = 0

C ===== coordonn,es factorielles =====
DO 70 IND = 1,NITOT1
READ (NFSAV)
(var1(I),I=1,NQEXA1),POIDS,(kidil(I),I=1,NID11)
DO 50 IAX = 1,NAXE
CRD = 0.
IF (valpr(IAX) .EQ. 0.) GOTO 50
DO 40 IC = 1,NCACT1
CRD = CRD + ctn(IC) * ccon(IC,IAx)
40 cind(IAX) = CRD
50 C ===== distance a l'origine =====
DORIG = 0.
DO 60 IC = 1,NCACT1
DORIG = DORIG + ctn(IC) * ctn(IC)
60 DOMAX = AMAX1(DOMAX,DORIG)

C ===== sauvegarde sur NFOUT =====
WRITE (NFOUT) (cind(I),I=1,NAXE),POIDS,DORIG,
(kidil(I),I=1,NID11),IZERO
70 CONTINUE

RETURN
END

SUBROUTINE DICG2
(NIACT1,NCACT1,PIACT1,VAR,NQEXA1,CTN,DORI,
- NFICH,VMOD)

INTEGER NIACT1, NCACT1, NQEXA1, NFICH
INTEGER LEC,ILLEC, IMP, NXCPA, NXLPA, NLIMP, NDICZ,
NDONZ,
- NDICA,NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR,
NDIC1,
- NDON1, NGUS1, NDON, NDIC, NDICB, NDONB, NDONA
REAL IFIN, IDEB, NOFF, I, IND, IC
REAL VAR(NQEXA1),CTN(NCACT1),VMOD(1,NCACT1)
REAL PIACT1, DORI, POIDS, PREL

COMMON /ENSOR/ LEC,ILLEC, IMP,NXCPA,NXLPA,NLIMP
COMMON /NUFIC/ NDICZ,NDONZ, NDICA,NDONA, NDICB,NDONB,
NDIC,NDON,
- NGUS,NGRI,NGRO, NSAV,NBAND,NBFOR
COMMON /NGREF/ NDIC1, NDON1, NGUS1

C ===== CODE COMME A DIMO2 =====
DORI=0.
IFIN = 0
C ===== CALCUL DES CENTRES DE GRAVITE PER TRANCHES =====
250 IDEB = IFIN + 1
NOFF = IFIN
IFIN = MINO (IFIN+NCACT1,NCACT1)
DO 260 I=1,NCACT1
VMOD(1,I)=0.
260 CONTINUE
REWIND NFICH
DO 270 IND = 1,NIACT1
READ (NFICH) (VAR(I),I=1,NQEXA1),POIDS
PREL = POIDS / PIACT1
DO 270 IC = IDEB,IFIN
VMOD(1,IC-NOFF) = VMOD(1,IC-NOFF) + PREL * CTN(IC)
270 CONTINUE
DO 280 IC = IDEB,IFIN
280 DORI = DORI + VMOD(1,IC-NOFF) * VMOD(1,IC-NOFF)
IF (IFIN .LT. NCACT1) GOTO 250

RETURN
END

SUBROUTINE STBCO
(EVAR,PTOT,CMOY,CVARAX,VMIN,VMAX,NCON,NAXE,
- CIND,NITOT1,NIACT1,LCORR,PNU1,NFGUS,NFOUT)

INTEGER NCON, NAXE, NITOT1, NIACT1, NFGUS, NFOUT, LCORR
INTEGER LEC,ILLEC, IMP, NXCPA, NXLPA, NLIMP, NDICZ,
NDONZ,
- NDICA,NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR,
NDIC1,
- NDON1, NGUS1, NGUS2, MDICZ,
MDONZ,NDICB,NDIC,NDONB,
- NDONA,MGUSZ, MGUSG, NMACZ, NMACA, NTRA, NTRA2,
NBASE

INTEGER I, IND, IC, NDON
REAL PTOT(NCON),CMOY(NCON),EVAR(NCON),
- VMIN(NCON),VMAX(NCON),CVARAX(NAXE),CIND(NAXE)
REAL PNU1, ZERO, RMIN, RMAX
REAL POIDS, VAL, PPRE, COEFM, COEFV,DELTC

COMMON /SEUIL/ ZERO,RMIN, RMAX
COMMON /NUFIC/ NDICZ,NDONZ, NDICA,NDONA, NDICB,NDONB,
NDIC,NDON,
- NGUS,NGRI,NGRO, NSAV,NBAND,NBFOR

```

```

COMMON /ENSOR/ LEC,ILLEC, IMP,NXCPA,NXLPA,NLIMP
COMMON /NGREF/ NDIC1, NDON1, NGUS1
COMMON /NGUS2/ NGUS2
common /nofic/ mdicz,mdonz,mgusg,mgusg, nmacz,nmaca,
- ntra,ntra2, nbase

Cini_staba===== INITIALISATIONS =====
DO 121 I = 1,NAXE
EVAR(I) = 0.
PTOT(I) = 0.
CMOY(I) = 0.
CVARAX(I) = 0.
VMIN(I) = RMAX
121 VMAX(I) = -RMAX
C ===== PREMIERE LECTURE : STATISTIQUES DE BASE =====
CALL OFICH(3,NFGUS)
DO 122 IND = 1,NIACT1
READ (NFGUS) (cind(I),I=1,NAXE),POIDS
DO 123 IC = 1,NAXE
VAL = cind(IC)
C ===== EFFECTIF ET POIDS DE LA VARIABLE =====
EVAR(IC) = EVAR(IC) + 1.
PPRE = PTOT(IC)
PTOT(IC) = PPRE + POIDS
IF (PTOT(IC) .LE. PNU1) GOTO 123
C ===== MOYENNE ET VARIANCE =====
COEFM = POIDS / PTOT(IC)
COEFV = PPRE * COEFM
DELTC = VAL - CMOY(IC)
CMOY(IC) = CMOY(IC) + COEFM * DELTC
CVARAX(IC) = CVARAX(IC) + COEFV * DELTC * DELTC
C ===== EXTREMA =====
VMIN(IC) = AMIN1 (VMIN(IC),VAL)
VMAX(IC) = AMAX1 (VMAX(IC),VAL)
123 CONTINUE
122 CONTINUE
C ===== PASSAGE AUX ECARTS-TYPES =====
DO 124 IC = 1,NAXE
IF (PTOT(IC) .GT. PNU1) THEN
IF (CVARAX(IC) .GT. 0.) THEN
CVARAX(IC) = SQRT (CVARAX(IC) / PTOT(IC))
ELSE
write(imp,*) 'Staba: +++ VARIABLE CONSTANTE +++'
ENDIF
ELSE
write(imp,*) 'Staba: +++ POIDS NUL +++'
ENDIF
124 CONTINUE

REWIND NFOUT
CALL OFICH(3,NFGUS)
DO 125 IND = 1,NITOT1
READ (NFGUS) (cind(I),I=1,NAXE),POIDS
C ===== RECODAGE, CENTRAGE, REDUCTION =====
DO 126 IC = 1,NAXE
cind(IC) = cind(IC) - CMOY(IC)
126 CONTINUE

C ===== ENREGISTREMENT DES DONNEES TRANSFORMEES =====
WRITE (NFOUT) (cind(I),I=1,NAXE),POIDS
125 CONTINUE

C == coord. centr,es+reduites -> NFOUT

RETURN
END

SUBROUTINE TESTCO(vect,vimp,corr0,kidv, NCACT1,
- NIACT,NIACT1,jorval,
- NFICOR0,NPAS0,NFICOR,NPAS)

INTEGER NCACT1, NIACT, NIACT1, NFICOR0, NPAS0, NFICOR,
NPAS
INTEGER NDICZ, NDONZ, NDICA, NDONA, NDICB, NDONB, NDIC,
NDON,
- NGUS, NGRI, NGRO, NSAV, NBAND, NBFOR, LEC, ILLEC,
IMP, NXCPA, NXLPA, NLIMP, NDIC1, NDON1, NGUS1,
NGUS2,
- MDICZ, MDONZ, MGUSZ, MGUSG, NMACZ, NMACA, NTRA,
NTRA2,
- NBASE
INTEGER I,J, II
INTEGER jorval(NCACT1)
CHARACTER*4 kidv(NCACT1)
REAL vect(NCACT1,NCACT1),corr0(NCACT1),vimp(NCACT1)
REAL ZA,ZB,ZC, ZD, ZE, ZF, Z01, Z02, Z03, Z04, Z11, Z12,
Z13,
- Z14, Z, PROB, PROBN, VDI, CORR

COMMON /NUFIC/ NDICZ,NDONZ, NDICA,NDONA, NDICB,NDONB,
NDIC,NDON,
- NGUS,NGRI,NGRO, NSAV,NBAND,NBFOR
COMMON /ENSOR/ LEC,ILLEC, IMP,NXCPA,NXLPA,NLIMP
common /ngref/ ndic1, ndon1, ngus1
common /ngus2/ ngus2
common /nofic/ mdicz,mdonz,mgusg,mgusg, nmacz,nmaca,

```

```

-          ntra,ntra2, nbase

write(imp,*) '**** TESTCORR ****'

DO 100 I=1,NCACT1
  corr0(I)=0
  DO 100 J=1,NCACT1
100    vect(I,J)=0

CALL OFICH (NPAS0,NFICOR0)
CALL OFICH (NPAS,NFICOR)
DO 200 I=1,NCACT1
  READ(NFICOR0) (corr0(II),II=1,NCACT1)
  DO 200 J=1,I-1
  DO 200 J=1,NCACT1
    READ(NFICOR) corr
    vect(I,J) = corr - corr0(J)
200 CONTINUE
WRITE(imp,2000)
NLIMP=1
WRITE(imp,3000)
NLIMP=NLIMP+4
CALL EDCOR(vect,kidv,vimp,NCACT1,1)

ZA = NIACT-3
ZB = NIACT1-3
ZC = 1 / ZA
ZD = 1 / ZB
ZE = ZC + ZD
ZZ = SQRT(ZE)
DO 250 I=1,NCACT1
  DO 250 J=1,NCACT1
250    vect(I,J)=0
CALL OFICH (NPAS0,NFICOR0)
CALL OFICH (NPAS,NFICOR)
DO 300 I=1,NCACT1
  READ(NFICOR0) (corr0(II),II=1,NCACT1)
  DO 300 J=1,I-1
  DO 300 J=1,NCACT1
    READ(NFICOR) corr
    IF (J .LT. I) THEN
      vdif = corr - corr0(J)
      Z01 = 1+corr0(J)
      Z02 = 1-corr0(J)

      Z03 = Z01 / Z02
      Z04 = ALOG(Z03) / 2
      Z11 = 1+corr
      Z12 = 1-corr
      Z13 = Z11 / Z12
      Z14 = ALOG(Z13) / 2
      Z=Z14-Z04
      IF (Z04 .GE. Z14) Z=Z04-Z14
      IF (Z04 .LT. Z14) Z=Z14-Z04
      vect(I,J) = Z/ZZ
    ENDIF
300 CONTINUE
WRITE(imp,2000)
NLIMP=1
WRITE(imp,4000)
NLIMP=NLIMP+4
CALL EDCOR(vect,kidv,vimp,NCACT1,1)

c===== impression des resultats en tableau =====
c    CALL EDTAB(NQACT,kidv,x2,ndlib,vtest,prob)
NLIMP = NLIMP + 4
WRITE(imp,5000)
WRITE(imp,5100)
CALL OFICH (NPAS0,NFICOR0)
CALL OFICH (NPAS,NFICOR)
DO 400 I=1,NCACT1
  READ(NFICOR0) (corr0(II),II=1,NCACT1)
  DO 400 J=1,I-1
  DO 400 J=1,NCACT1
    READ(NFICOR) corr
    vdif = corr-corr0(J)
    IF (J .LT. I) THEN
      prob = PROBN(vect(I,J),1)
C = PROBLEMA
      WRITE(imp,5200) jorval(I),kidv(I),jorval(J),
-          kidv(J),corr,corr0(J),vdif,vect(I,J),prob
C = PROBLEMA
    ENDIF
400 CONTINUE

2000 FORMAT (1H1)
3000 FORMAT (3H 0/39H TABLEAU DU DIFFERENCES DE CORRELATIONS/
-          3X,39(1H-))
4000 FORMAT (3H 0/36H TABLEAU VTEST POUR LES CORRELATIONS/
-          3X,36(1H-))

5000 FORMAT (1H1,1H+,74(1H-),1H+/2H |,3X,17H VALEURS-TEST DES,
-          26H MATRICES DES CORRELATIONS,27X,2H |/2H
+ ,74(1H-),1H+)
5100 FORMAT (2H |,12H VARIABLE 1,5X,11H VARIABLE 2,7X,6H
corr1,
-          7H corr2,10H dif_corr,8H v_test,6H prob,3H
|/2H +,
-          74(1H-),1H+)
5200 FORMAT (3X,15,'-',A4,6X,I4,'-',A4,6X,

```

```

-          5P8.2,1H1/2H +,74(1H-),1H+)

RETURN
END

SUBROUTINE TRASP(XMAT,IDIM1,IDIM2,XMATT)
  INTEGER IDIM1, IDIM2, IXTT, JXTT
  REAL XMAT(IDIM1, IDIM2), XMATT(IDIM2, IDIM1)
C    COMMON /ENSOR/ LEC,ILLEC, IMP,NXCPA,NXLPA,NLIMP

DO 100 IXTT=1, IDIM2
  DO 100 JXTT=1, IDIM1
    XMATT(IXTT,JXTT)=XMAT(JXTT,IXTT)
100 CONTINUE

RETURN
END

SUBROUTINE
SIGNE2(XMAT0, IDIMIO, XMAT1, IDIMI1, IDIMJ, IDIM, XMATT, X)
  INTEGER IDIMIO, IDIMI1, IDIMJ, IDIM
  INTEGER LEC, ILLEC, IMP, NXCPA, NXLPA, NLIMP
  INTEGER NDO, I, J
  REAL XMAT0(IDIMIO, IDIMJ), XMAT1(IDIMI1, IDIMJ),
-          XMATT(IDIMJ, IDIMI1), X(IDIMIO, IDIMI1)
  COMMON /ENSOR/ LEC, ILLEC, IMP, NXCPA, NXLPA, NLIMP

  CALL TRASP(XMAT1, IDIMI1, IDIMJ, XMATT)
  CALL MULTM(XMAT0, IDIMIO, IDIMJ, XMATT, IDIMJ, IDIMI1, IDIMJ, X)

  NDO=0
  === ON VERIFIE LE SIGNE SUELEMENT PAR LES NDIM PREMIERS
  DO 200 I=1, IDIM
    IF (X(I,I) .EQ. 0.) NDO=NDO+1
    IF (X(I,I) .LT. 0.) THEN
      DO 250 J=1, IDIMJ
        XMAT1(I,J)=-XMAT1(I,J)
250      CONTINUE
    ENDIF
200 CONTINUE

RETURN
END

SUBROUTINE MULTM(XMAT1, IDIMI1, IDIM12, XMAT2, IDIM21, IDIM22,
-          IDIMC, XMAT3)
  INTEGER IDIMI1, IDIM12, IDIM21, IDIM22, IDIMC
  INTEGER I, J, K
  INTEGER LEC, ILLEC, IMP, NXCPA, NXLPA, NLIMP
  REAL XMAT1(IDIMI1, IDIM12), XMAT2(IDIM21, IDIM22),
-          XMAT3(IDIMI1, IDIM22)
  COMMON /ENSOR/ LEC, ILLEC, IMP, NXCPA, NXLPA, NLIMP

DO 50 I=1, IDIMI1
  DO 50 J=1, IDIM22
    XMAT3(I,J)=0.
50 CONTINUE

DO 100 I=1, IDIMI1
  DO 100 J=1, IDIM22
    DO 150 K=1, IDIMC
      XMAT3(I,J)=XMAT3(I,J)+XMAT1(I,K)*XMAT2(K,J)
150    CONTINUE
c    write(imp,*) 'multm. i,j,valor: ',i,j,xmat3(i,j)
100 CONTINUE
RETURN
END

```

La resta de subrutines (ECCIF, EDCDG, CORMO2, DIMO2, STABA, COREL, ECCVA) són auxiliars a la rutina principal

Annex3. Resultats de la inserció per a 3.7

Annex3.1 Subconjunt mínim de variables comunes (MIVAR)

```

RECONSTRUCTION DES 1 PREMIERS AXES
CRITERE INITIAUX AVEC 3 VARIABLES: 2.007
(somme des correlations carrees des premiers axes ponderees par les valeurs propres)
VALEURS PROPRES DES 1 PREMIERS AXES:
2.0074744
LISTE DES 3 VARIABLES INITIAUX:
1:C1 j1
2:C2 j2
3:C3 j3
ELIMINATION DES VARIABLES PAS A PAS JUSQUA 1
Step Var. eliminee Critere
3-> 2 VARS. 3 1.923 *****
2-> 1 VARS. 1 0.965 *****
INCORPORATION DES VARIABLES PAS A PAS JUSQUA 3
Step Var. ajoutee Critere
1-> 2 VARS. 2 1.931 *****
2-> 3 VARS. 1 2.007 *****

```

Annex3.2 Estadístiques descriptives de les variables contínues comunes

STATISTIQUES SOMMAIRES DES VARIABLES CONTINUES						
EFFECTIF TOTAL : 80			POIDS TOTAL : 80.00			
NUM . IDEN - LIBELLE	EFFECTIF	POIDS	MOYENNE	ECART-TYPE	MINIMUM	MAXIMUM
1 . C1 - j1	80	80.00	0.00	3.50	-7.38	7.39
2 . C2 - j2	80	80.00	0.00	3.42	-7.27	7.68
3 . C3 - j3	80	80.00	0.00	0.01	-0.02	0.03

STATISTIQUES SOMMAIRES DES VARIABLES CONTINUES						
EFFECTIF TOTAL : 50			POIDS TOTAL : 50.00			
NUM . IDEN - LIBELLE	EFFECTIF	POIDS	MOYENNE	ECART-TYPE	MINIMUM	MAXIMUM
1 . C1 - j1	50	50.00	2.05	2.86	-3.73	7.53
2 . C2 - j2	50	50.00	-0.08	3.14	-6.20	5.66
3 . C3 - j3	50	50.00	0.00	0.01	-0.02	0.03

Annex3.3 Test de correlacions

TABLEAU DU DIFFERENCES DE CORRELATIONS

	C1	C2	C3
C1	0.00		
C2	0.05	0.00	
C3	-0.18	0.19	0.00

TABLEAU VTEST POUR LES CORRELATIONS

	C1	C2	C3
C1	0.00		
C2	2.51	0.00	
C3	-1.00	1.02	0.00

VALEURS-TEST DES MATRICES DES CORRELATIONS						
VARIABLE 1	VARIABLE 2	corr1	corr2	dif_corr	v_test	prob
2. C2	1. C1	-0.91	-0.96	0.05	2.51	0.01
3. C3	1. C1	-0.05	0.14	-0.18	-1.00	0.84
3. C3	2. C2	0.03	-0.16	0.19	1.02	0.15

Annex3.4 Coordenades de les variables contínues puntuals

COORDONNEES DES VARIABLES SUR LES AXES 1 A 3
VARIABLES ACTIVES

VARIABLES		COORDONNEES					CORRELATIONS VARIABLE-FACTEUR					ANCIENS AXES UNITAIRES			
IDEN - LIBELLE COURT		1	2	3	0	0	1	2	3	0	0	1	2	3	0
C1 - j1		0.96	-0.30	-0.06	0.00	0.00	0.96	-0.30	-0.06	0.00	0.00	0.00	0.00	0.00	0.00
C2 - j2		-0.97	0.27	0.47	0.00	0.00	-0.97	0.27	0.47	0.00	0.00	0.00	0.00	0.00	0.00
C3 - j3		0.12	0.97	0.03	0.00	0.00	0.12	0.97	0.03	0.00	0.00	0.00	0.00	0.00	0.00

Annex3.5 Coordenades dels individus puntuals

COORDONNEES, CONTRIBUTIONS ET COSINUS CARRES DES INDIVIDUS SUR LES AXES 1 A 3

INDIVIDUS			COORDONNEES					CONTRIBUTIONS					COSINUS CARRES				
IDENTIFICATEUR	P.REL	DISTO	1	2	3	0	0	1	2	3	0	0	1	2	3	0	0
01	2.00	5.39	-2.06	-0.98	0.43	0.00	0.00	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00
02	2.00	1.07	-0.91	-0.28	0.41	0.00	0.00	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00
03	2.00	0.61	-0.40	-0.67	0.00	0.00	0.00	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00
04	2.00	2.16	-1.23	-0.43	0.68	0.00	0.00	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00
05	2.00	2.39	0.89	1.15	0.52	0.00	0.00	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00
	...																
48	2.00	1.17	0.85	0.35	0.58	0.00	0.00	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00
49	2.00	2.03	-0.97	-0.29	1.01	0.00	0.00	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00
50	2.00	8.59	1.93	-2.19	-0.25	0.00	0.00	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00

Annex3.6 Coordenades de l'individu “centre de gravetat”

COORDONNEES ET VALEURS-TEST DU CENTRE GRAVIT,E SUR LES AXES 1 A 3

VALEURS-TEST						COORDONNEES					DISTO
1	2	3	0	0		1	2	3	0	0	
Centre	2.7	-2.1	2.8	0.0	0.0	0.38	-0.30	0.39	0.00	0.00	0.38

Annex3.7 Coordenades de la variable categòrica

COORDONNEES ET VALEURS-TEST DES MODALITES
AXES 1 A 3

MODALITES			VALEURS-TEST					COORDONNEES					
IDEN - LIBELLE	EFF.	P.ABS	1	2	3	0	0	1	2	3	0	0	DISTO.
4 . M													
AA_1 - 2=1	23	23.00	-1.9	-1.1	9.7	0.0	0.0	-0.40	-0.17	0.45	0.00	0.00	0.39
AA_2 - 2=2	27	27.00	5.7	-3.1	8.6	0.0	0.0	1.04	-0.41	0.34	0.00	0.00	1.37

Annex4. Resultats de la inserció per a 4

Annex4.1 Subconjunt mínim de variables comunes (MIVAR)

```
RECONSTRUCTION DES 2 PREMIERS AXES
CRITERE INITIAL AVEC 36 VARIABLES: 11.589
(somme des correlations carrees des premiers axes ponderees par les valeurs propres)
VALEURS PROPRES DES 2 PREMIERS AXES:
7.3740115
4.2153749
LISTE DES 36 VARIABLES INITIAUX:
1:C18 t11
2:C19 t12
3:C20 t13
...
35:C64 c315
36:C65 c316
ELIMINATION DES VARIABLES PAS A PAS JUSQUA 12
Step Var. eliminee Critere
36 -> 35 VARS. 32 11.588 *****
35 -> 34 VARS. 13 11.583 *****
34 -> 33 VARS. 7 11.576 *****
33 -> 32 VARS. 31 11.565 *****
32 -> 31 VARS. 8 11.562 *****
31 -> 30 VARS. 33 11.536 *****
30 -> 29 VARS. 19 11.521 *****
29 -> 28 VARS. 25 11.503 *****
28 -> 27 VARS. 1 11.482 *****
27 -> 26 VARS. 2 11.453 *****
26 -> 25 VARS. 26 11.418 *****
25 -> 24 VARS. 35 11.379 *****
24 -> 23 VARS. 34 11.339 *****
23 -> 22 VARS. 14 11.294 *****
22 -> 21 VARS. 3 11.242 *****
21 -> 20 VARS. 20 11.185 *****
20 -> 19 VARS. 9 11.110 *****
19 -> 18 VARS. 5 11.029 *****
18 -> 17 VARS. 11 10.938 *****
17 -> 16 VARS. 12 10.831 *****
16 -> 15 VARS. 36 10.724 *****
15 -> 14 VARS. 10 10.611 *****
14 -> 13 VARS. 4 10.484 *****
13 -> 12 VARS. 21 10.302 *****
INCORPORATION DES VARIABLES PAS A PAS JUSQUA 36
Step Var. ajoutee Critere
12 -> 13 VARS. 21 10.484 *****
13 -> 14 VARS. 4 10.611 *****
14 -> 15 VARS. 10 10.724 *****
15 -> 16 VARS. 36 10.831 *****
16 -> 17 VARS. 12 10.938 *****
17 -> 18 VARS. 11 11.029 *****
18 -> 19 VARS. 5 11.110 *****
19 -> 20 VARS. 9 11.185 *****
20 -> 21 VARS. 20 11.242 *****
21 -> 22 VARS. 3 11.294 *****
22 -> 23 VARS. 14 11.339 *****
23 -> 24 VARS. 34 11.379 *****
24 -> 25 VARS. 35 11.418 *****
25 -> 26 VARS. 26 11.453 *****
26 -> 27 VARS. 2 11.482 *****
27 -> 28 VARS. 1 11.503 *****
28 -> 29 VARS. 25 11.521 *****
29 -> 30 VARS. 19 11.536 *****
30 -> 31 VARS. 33 11.552 *****
31 -> 32 VARS. 8 11.565 *****
32 -> 33 VARS. 31 11.576 *****
33 -> 34 VARS. 7 11.583 *****
34 -> 35 VARS. 13 11.588 *****
35 -> 36 VARS. 32 11.589 *****
```

Annex4.2 Estadístiques descriptives de les variables contínues comunes

STATISTIQUES SOMMAIRES DES VARIABLES CONTINUES

EFFECTIF TOTAL : 380 POIDS TOTAL : 380.00

NUM . IDEN - LIBELLE	EFFECTIF	POIDS	MOYENNE	ECART-TYPE	MINIMUM	MAXIMUM
22 . C23 - t116	380	380.00	3.68	1.23	1.00	5.00
31 . C32 - a313	380	380.00	4.58	1.00	1.00	5.00
32 . C33 - a314	380	380.00	4.18	1.28	1.00	5.00
33 . C34 - a315	380	380.00	4.02	1.28	1.00	5.00
34 . C35 - a316	380	380.00	3.47	1.21	1.00	5.00
38 . C39 - t514	380	380.00	3.97	1.40	1.00	5.00
39 . C40 - t515	380	380.00	4.02	1.30	1.00	5.00
40 . C41 - t516	380	380.00	3.42	1.27	1.00	5.00
55 . C56 - v313	380	380.00	4.47	1.17	1.00	5.00
56 . C57 - v314	380	380.00	3.61	1.67	1.00	5.00
57 . C58 - v315	380	380.00	4.18	1.23	1.00	5.00
58 . C59 - v316	380	380.00	3.65	1.37	1.00	5.00

b2STATISTIQUES SOMMAIRES DES VARIABLES CONTINUES

EFFECTIF TOTAL : 837 POIDS TOTAL : 837.00

NUM . IDEN - LIBELLE	EFFECTIF	POIDS	MOYENNE	ECART-TYPE	MINIMUM	MAXIMUM
21 . C22 - t116	837	837.00	3.61	1.13	1.00	5.00
30 . C31 - a313	837	837.00	4.65	0.88	1.00	5.00
31 . C32 - a314	837	837.00	4.02	1.36	1.00	5.00
32 . C33 - a315	837	837.00	4.17	1.16	1.00	5.00
33 . C34 - a316	837	837.00	3.42	1.14	1.00	5.00
37 . C38 - t514	837	837.00	4.25	1.19	1.00	5.00
38 . C39 - t515	837	837.00	4.15	1.17	1.00	5.00
39 . C40 - t516	837	837.00	3.57	1.09	1.00	5.00
54 . C55 - v313	837	837.00	4.47	1.18	1.00	5.00
55 . C56 - v314	837	837.00	3.62	1.68	1.00	5.00
56 . C57 - v315	837	837.00	4.24	1.22	1.00	5.00
57 . C58 - v316	837	837.00	3.74	1.37	1.00	5.00
250 . C251 - suavitzans abans	837	837.00	266.66	461.00	-1.00	5832.00
251 . C252 - flor abans	837	837.00	38.05	174.55	-1.00	1963.00
252 . C253 - suavitzants durant	837	837.00	264.00	425.21	-1.00	5037.00
253 . C254 - flor durant	837	837.00	39.44	160.37	-1.00	1380.00
254 . C255 - caldos abans	837	837.00	78.00	214.25	-1.00	3600.00
255 . C256 - starlux abans	837	837.00	23.29	118.03	-1.00	1729.00
256 . C257 - caldos durant	837	837.00	101.61	191.63	-1.00	1683.00
257 . C258 - starlux durant	837	837.00	34.04	114.36	-1.00	953.00

Annex4.3 Correlacions entre els eixos

CORRELATIONS ENTRE AXES DU NGUS ET NGUS1

AXES du NGUS1										
0.91	-0.36	-0.13	-0.01	-0.02	0.09	0.09	-0.07	-0.07	-0.03	
0.38	0.92	0.07	0.03	-0.01	-0.01	-0.06	-0.06	0.04	0.05	
0.05	-0.06	0.14	0.33	0.59	-0.56	-0.21	-0.31	-0.09	-0.20	
-0.09	0.03	-0.63	0.12	-0.27	-0.04	-0.28	-0.30	0.16	-0.20	
-0.03	0.07	-0.38	-0.61	0.17	-0.32	0.15	-0.06	-0.44	0.05	
0.00	0.04	-0.20	-0.06	0.69	0.35	0.09	0.41	0.22	0.00	
0.02	0.06	-0.34	0.04	-0.07	-0.45	0.47	0.25	0.48	-0.04	
0.05	-0.02	0.24	0.08	-0.20	-0.41	0.10	0.40	-0.24	0.33	
0.06	-0.09	-0.04	0.35	-0.02	-0.14	0.00	0.02	0.30	0.31	
0.03	0.02	-0.37	0.25	0.00	-0.01	-0.53	0.48	-0.31	0.19	

Annex4.4 Coordenades de les variables contínues puntuals

COORDONNEES DES VARIABLES SUR LES AXES 1 A 2

VARIABLES ACTIVES

VARIABLES	COORDONNEES	CORRELATIONS VARIABLE-FACTEUR	ANCIENS AXES UNITAIRES
IDEN - LIBELLE COURT	1 2 0 0 0	1 2 0 0 0	1 2 0 0
C22 - t116	0.30 -0.13 0.00 0.00 0.00	0.26 -0.11 0.00 0.00 0.00	0.00 0.00 0.00 0.00 0.00
C31 - a313	0.23 -0.25 0.00 0.00 0.00	0.26 -0.29 0.00 0.00 0.00	0.00 0.00 0.00 0.00 0.00
C32 - a314	0.62 -0.52 0.00 0.00 0.00	0.46 -0.39 0.00 0.00 0.00	0.00 0.00 0.00 0.00 0.00
C33 - a315	0.45 -0.59 0.00 0.00 0.00	0.39 -0.51 0.00 0.00 0.00	0.00 0.00 0.00 0.00 0.00
C34 - a316	0.48 -0.47 0.00 0.00 0.00	0.42 -0.42 0.00 0.00 0.00	0.00 0.00 0.00 0.00 0.00
C38 - t514	0.48 -0.42 0.00 0.00 0.00	0.40 -0.36 0.00 0.00 0.00	0.00 0.00 0.00 0.00 0.00
C39 - t515	0.35 -0.60 0.00 0.00 0.00	0.30 -0.51 0.00 0.00 0.00	0.00 0.00 0.00 0.00 0.00
C40 - t516	0.44 -0.39 0.00 0.00 0.00	0.41 -0.36 0.00 0.00 0.00	0.00 0.00 0.00 0.00 0.00
C55 - v313	-0.45 -0.52 0.00 0.00 0.00	-0.38 -0.44 0.00 0.00 0.00	0.00 0.00 0.00 0.00 0.00
C56 - v314	-1.22 -1.00 0.00 0.00 0.00	-0.73 -0.60 0.00 0.00 0.00	0.00 0.00 0.00 0.00 0.00
C57 - v315	-0.59 -0.60 0.00 0.00 0.00	-0.48 -0.49 0.00 0.00 0.00	0.00 0.00 0.00 0.00 0.00
C58 - v316	-0.91 -0.58 0.00 0.00 0.00	-0.66 -0.42 0.00 0.00 0.00	0.00 0.00 0.00 0.00 0.00

VARIABLES ILLUSTRATIVES

VARIABLES	COORDONNEES					CORRELATIONS VARIABLE-FACTEUR				
DEN - LIBELLE COURT	1	2	0	0	0	1	2	0	0	0
C251 - suavitizans abans	-45.00	6.22	0.00	0.00	0.00	-0.10	0.01	0.00	0.00	0.00
C252 - flor abans	-4.39	0.50	0.00	0.00	0.00	-0.03	0.00	0.00	0.00	0.00
C253 - suavitizants durant	-7.23	25.54	0.00	0.00	0.00	-0.02	0.06	0.00	0.00	0.00
C254 - flor durant	-3.88	-4.47	0.00	0.00	0.00	-0.02	-0.03	0.00	0.00	0.00
C255 - caldos abans	-14.81	-1.75	0.00	0.00	0.00	-0.07	-0.01	0.00	0.00	0.00
C256 - starlux abans	-14.74	-6.20	0.00	0.00	0.00	-0.12	-0.05	0.00	0.00	0.00
C257 - caldos durant	-21.16	-2.87	0.00	0.00	0.00	-0.11	-0.01	0.00	0.00	0.00
C258 - starlux durant	-17.66	-12.38	0.00	0.00	0.00	-0.15	-0.11	0.00	0.00	0.00

Annex4.5 Coordenades de l'individu "centre de gravetat"

COORDONNEES ET VALEURS-TEST DU CENTRE GRAVIT,E SUR LES AXES 1 A 2

	VALEURS-TEST					COORDONNEES					DISTO
	1	2	0	0	0	1	2	0	0	0	
Centre	2.2	-6.0	0.0	0.0	0.0	0.07	-0.21	0.00	0.00	0.00	0.19

Annex4.6 Coordenades de les variables categòriques

COORDONNEES ET VALEURS-TEST DES MODALITES AXES 1 A 2

MODALITES			VALEURS-TEST					COORDONNEES						
DEN	LIBELLE	EFF.	P.ABS	1	2	0	0	0	1	2	0	0	0	DISTO.
6 . aspirador														
E_1	- C7=1	498	498.00	5.2	-3.6	0.0	0.0	0.0	0.31	-0.19	0.00	0.00	0.00	0.26
E_2	- C7=2	339	339.00	-3.1	-3.0	0.0	0.0	0.0	-0.28	-0.23	0.00	0.00	0.00	0.33
10 . automobil														
AI_1	- C11=1	171	171.00	0.7	-2.0	0.0	0.0	0.0	0.10	-0.25	0.00	0.00	0.00	0.29
AI_2	- C11=2	519	519.00	-1.0	-4.6	0.0	0.0	0.0	-0.06	-0.23	0.00	0.00	0.00	0.24
AI_3	- C11=3	131	131.00	3.3	-0.4	0.0	0.0	0.0	0.56	-0.07	0.00	0.00	0.00	0.46
AI_4	- C11=4	16	16.00	0.4	-0.7	0.0	0.0	0.0	0.23	-0.33	0.00	0.00	0.00	0.77

Bibliografia

Inserció de dades d'enquesta, Tesi Doctoral de Roser Rius, Barcelona 1998.

Introducción al análisis factorial, de M. Cuesta i F.J. Herrero, www.uniovi.es.

Análisis de Componentes Principales, de E. Barrull, 1992, www.biopsychology.org.

Resum seminari “Enquêtes de référence, greffe d'enquêtes et fusion entre fichiers d'enquêtes” per J.P. Pagès, J.M. Gautier i A. Iliakopoulos, 1991, www.inded.fr.

Introduction to Fortran 90, The parallel computer center, www.pcc.qub.ac.uk.

Migrating to Fortran 90, James F. Kerrigan, web.mse.uiuc.edu.